REPORT RESUMES

ED 010 389 64

THE ROLE OF STUDENT RESPONSE IN LEARNING FROM THE NEW EDUCATIONAL MEDIA.

BY- MAY, MARK A.
MAY, MARK A., HAMDEN, CONN.

REPORT NUMBER BR-5-0999 PUB DATE AUG 66
REPORT NUMBER NDEA-VIIB-530
CONTRACT OEC-5-16-006
EDRS PRICE MF-$0.36 HC-$8.00 200P.

DESCRIPTORS- *COVERT RESPONSE, *OVERT RESPONSE, *LEARNING PROCESSES, INSTRUCTIONAL MATERIALS, STUDENT CHARACTERISTICS, *BEHAVIOR, *LEARNING THEORIES, CINCINNATI, OHIO, HAMDEN, CONNECTICUT

COMBINATIONS OF OVERT AND COVERT RESPONSE PRACTICES WERE ANALYZED TO DETERMINE OPTIMAL COMBINATIONS FOR THE LEARNING PROCESS FOR (1) DIFFERENT LEARNING TASKS, (2) STUDENTS OF DIFFERENT ABILITIES, AND (3) DIFFERENT MEDIA OF PRESENTATION. IT WAS OBSERVED THAT PRACTICALLY ALL FORMS OF HUMAN LEARNING INVOLVE SUCH COVERT ACTIVITIES AS OBSERVING, LISTENING, READING, AND COGITATING, BUT DO NOT ALWAYS REQUIRE OVERT FORMS OF BEHAVIOR. THE PROBLEM WAS, THEREFORE, TO DETERMINE THE CONTRIBUTIONS TO LEARNING MADE BY OVERT RESPONDING, CONSIDERING THAT INTERFERENCE WITH COVERT PROCESSES WOULD SUBSTANTIALLY REDUCE LEARNING RATE. THE RESEARCH EFFORT WAS EXPLORATORY AND DIVERSE AS DATA WERE OBTAINED FROM APPROXIMATELY 50 SHORT EXPERIMENTS WHICH USED A WIDE VARIETY OF PROGRAMS, INSTRUCTIONAL MATERIALS, SAMPLE POPULATIONS, AND TEST METHODS. FROM THE RESEARCH OF APPROXIMATELY 100 INDIVIDUALS USING A VARIETY OF PROGRAMS AND INSTRUCTIONAL MATERIALS, A SET OF TENTATIVE PROPOSITIONS WAS SET FORTH--(1) IN AMOUNTS LEARNED PER UNIT OF TIME, COVERT RESPONDING IS MORE EFFICIENT THAN COVERT PLUS OVERT, SINCE OVERT RESPONDING ALWAYS REQUIRES ADDITIONAL TIME, (2) WHEN INSTRUCTION IS CONDUCTED IN THREE PHASES OF--(A) OBSERVING, LISTENING, OR STUDYING, (B) TESTING, ANSWERING, OR RECITING, AND (C) CONFIRMATION OR CORRECTION,--OVERT RESPONDING CONTRIBUTES MORE TO LEARNING IN THE SECOND PHASE THAN IT DOES IN EITHER THE FIRST OR THIRD, AND (3) NO ONE RESPONSE MODE IS OPTIMAL IN ALL SITUATIONS. SOME INDIVIDUALS, HOWEVER, ADVOCATE OVERT ANSWERING RESPONSES FOR ALL FORMS OF PROGRAMED INSTRUCTION. (JH)

ED010389

B-530

5-0999

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

THE ROLE OF STUDENT RESPONSE IN LEARNING FROM THE NEW EDUCATIONAL MEDIA

Mark A. May

THE ROLE OF STUDENT RESPONSE IN LEARNING FROM THE NEW EDUCATIONAL MEDIA

Mark A. May

U. S. Department of Health, Education, and Welfare

Office of Education

Contract No. OE-5-16-006

Final Report

August 1966

Foreword

This is the third in a series of working papers prepared for the Media Branch of the USOE. The first two dealt with stimulus variables. This one deals with response variables. The two sets of variables are treated separately only for the purposes of exposition.

In the paper on "Enhancements and Simplifications of AV Presentations," stimulus variables were considered from the standpoint of the functions they perform in relation to responses. The four main functions are motivation, reinforcement, cue identification, and simplification. In the paper on "Picture-Word Relationships in AV Presentations," various combinations of stimuli such as visual-verbal, auditory verbal, visual and nonverbal were considered from the standpoint of their relative merits for motivating, eliciting, and reinforcing the responses that are required by different learning tasks.

Three main types of responses were considered: preparatory, acquisition, and consolidating.[1] Preparatory responses include: (a) sensory orientation - responses of adjusting the eyes and ears to a presentation; (b) paying attention responses - putting one's mind on the presentation; (c) targeting responses - selective attention to the relevant cues and ignoring the irrelevant ones; and (d) perceptual responses - perceiving the correct meanings of the crucial cues.

---

[1] This threefold classification of response is in some respects parallel to, but, in other respects different from, the one proposed by Gagne and Bolles (1959) who described two major sets of factors called "readiness" and "associative."

The first two of these four classes of preparatory responses are essential for all learning regardless of the learning task. The responses of concentrating attention on the crucial cues and of making the correct perceptual responses to these cues are defined by the educational objectives, specified by the learning task, and illustrated by properly constructed criterion tests.

Included in the category of acquisition responses are all of the new responses to be learned from a presentation. They differ from preparatory responses in that they are assumed not to have been previously learned. In most presentations a certain amount of the material may already have been learned, as indicated by scores on pretest. The new responses that are acquired are defined operationally as gains from pre- to post tests.

The kinds of new responses to be acquired are determined, to a large degree, by the rature of the learning tasks. In one of the previous papers cited above (151), learning tasks are put into two main categories - reproductive and productive-constructive, as suggested by Gagne (64). Reproductive tasks are memorizing lists of names, verses of poetry, the vocabulary of a foreign language, or the movements made in a demonstration of a skilled performance.

Productive-constructive tasks are those that require the learner to give in his own words the "substance" of what was learned, as in essay examinations, or to utilize what has been learned in solving a problem or performing successfully in related but different situations or under different conditions. Learning to improvise, to invent, to cope with a problem situation may also be included in this broad category of tasks.

The third major category is "consolidating responses" which are defined operationally by amounts retained after varying periods of delay. This is a separate category because, in all experiments in which tests of immediate and delayed recall have been used, there is a drop or loss from the immediate to the delayed test. This is the familiar phenomenon of forgetting. From the standpoint of educational objectives, the important thing is not how much a student learned from a presentation or lesson, but how much of what was learned is retained and can be recalled and used. There is ample evidence that a great deal of what is learned in school is forgotten and often to the point of no return or recall. The conditions on which learning and retention depend are to some extent the same, but there are other conditions affecting retention. One is the degree of overlearning; another is the extent to which the learned response is continuously used, applied, and practiced.

The responses in each of the above three major categories are mainly covert. This is particularly true in learning from AV presentations that make no provision for active student participation or practice during the learning sessions. The preparatory responses of orienting the eyes and ears to the screen and those of looking squarely at the crucial visual cues, are mainly overt. But the responses of perceiving the correct meanings of the relevant cues, and those of associating them with the to-be-learned responses, and those of reinforcing and consolidating these associations are covert. That such covert responding is effective is demonstrated by the gains in knowledge from pretests to post tests.

The main purpose of the present paper is to present evidence on the extent to which gains in knowledge, or in skills, are enhanced by making provision during presentations for overt responding and active student participation or practice, particularly in connection with responses of acquisition and consolidation. The experimental literature on student participation in learning from films has been reviewed by Allen (1) under five main headings: (1) verbalization of response (2) perceptual-motor responses (3) knowledge of results (4) mental practice, and (5) note-taking. Allen's review is extended by Lumsdaine (130), who added the variables of time, amount of active response, direct-practice effects vs. side effects such as motivation, form of overt response in verbal learning, overt and covert responding, feedback, reinforcement and knowledge of results, guidance, cueing or prompting, interaction of prompting and overt responses, organizational and sequencing factors, size of step and self-pacing practice. Lumsdaine's review is condensed somewhat and a few additional experiments are included.in a more recent review by Lumsdaine and May (135).

Other reviews of the experimental literature on programmed instruction are those by Goldstein and Gotkin (81), Silberman (194), Briggs and Angell (17), Briggs and Hamilton (18), Champeau (28), and Holland (98). An extensive annotated bibliography of research on programmed insturction has been published by Schramm (185).

In addition to reviews of research literature, a number of collections of papers have been compiled and edited by: Galanter (72), Lumsdaine and Glaser (136), Coulson (37), Gage (63), Decceco (45), and Glaser (77).

During the decade 1956-1966 no less than 50 experimental comparisons were made between alternate modes of responding to audio-visual presentations and to printed instructional programs. In addition, as many as 150 experiments have been reported that bear directly or indirectly on optimal response modes.

It is not the purpose of this paper to re-review all of this literature but rather to select from it experiments and theoretical discussions that relate to modes of responding which have been found to be the most effective and efficient for the accomplishment of various kinds of learning tasks, by students of varying abilities.

I. INTRODUCTION

1. The Problem of Student Response.

Proceeding from the basic premise that all learning is by doing, the problem is to discover the kinds of learning activities that are required for the achievement of educational objectives whatever they may be. It is a problem of practical importance for teachers. Psychologists and some educators are insisting, more than ever before, that all educational objectives be defined in terms of to-be-achieved abilities and capabilities. What is to be taught, and how, is determined by behavioral changes to be sought. When objectives are so specified, all learning activities are directed toward clearly defined goals and become central in the instructional process.

The human organism is capable of making and learning to make a wide variety of responses to environmental and internal stimuli. These responses are so diverse and vary along so many different dimensions as to defy detailed or satisfactory classification. It is possible and useful, however, to group them into (a) muscular responses which are either directly observable or measureable with appropriate instruments, and (b) those which are inferred from observable antecedent events and changes in behavior. The former are commonly called overt and the latter covert responses. Included in the category of covert responses are the so-called "higher mental processes" of perception, cognition, memory, thinking, reasoning, purpose, incentives, foresight, etc.

The problem is to discover the relative contribution to learning made by such covert activities as observing, listening, reading, studying and cogitating on the one hand, and overt responding such as reciting, answering questions, engaging in discussions, on the other. Both modes of responding have their advantages and disadvantages for learning. The relation between these two phases of the instructional process is illustrated by the classical experiment of Gates on relative amount of time spent in study and recitation (73).

The instructional process is viewed here as consisting of two and often three distinguishable phases. As indicated above, the first phase is that of observing, listening, reading and studying during which the responses are mostly covert, although they may be overt as in reading aloud. The second phase is one of reciting, answering questions, taking tests, during which the responses are usually overt but may also be only covert. The third phase, prominent in programmed instruction, is one of confirmation, correction, reinforcement and informational "feedback."

Assuming that the responses in the first phase are mainly covert except for eye movement and incipient lip movements, the question arises as to what contributions are made to learning by introducing the second and third phases, and particularly what contributions are made by overt responding to the stimuli presented in these two phases. The problem is not one of overt versus covert responding, as though they were independent of each other, but rather that of discovering the contribution to learning made by overt responding in addition to that inevitably made by the covert responses of attention, perception, memory, and cognition.

Despite its many practical and theoretical advantages, which are listed below, overt responding is not a necessary condition of all forms of learning. If it were, then as Lumsdaine has pointed out (130), "No one would gain anything from reading a book without reading it aloud, from watching a film without mimicking its actions, or from listening to a lecture without chorusing, echoing responses" (p. 610). The essential condition is that some form of the responses that are to be learned must occur during the course of training and in the presence of relevant cues. These responses may be either covert alone, or covert followed by overt, or vice versa.

There are, however, some learning tasks for which active overt practice is essential for attaining a high level of skill. It is quite unlikely that anyone could become a fast typist, an accomplished pianist, fluent in speaking a foreign language, or a champion athlete without considerable amount of active practice. The problem is to discover the combination of overt and covert practice that is optimal for different learning tasks, for students of different abilities, and for different media of presentation.

If overt responding is not an essential condition for the learning of all tasks, under all conditions, and for all levels of student ability, the question arises as to whether or not some form of covert responding is a universally essential condition of human learning. That verbal learning does occur by reading and by observation, as well as by "thinking" the correct answers to frames in a programmed lesson, is a well established fact. There is also evidence that when covert responding is interfered with, the rate of learning is substantially reduced. For example, Kendler, Kendler, and Cook (110) tried to prevent covert responding by filling the period allowed for responding with noninformative verbiage. They succeeded in cutting down substantially the amount learned when compared with a control group. Thomson (217) required some of her subjects to count backward by two's from one hundred to zero while watching the assembly of the parts of a wooden puzzle. Again the amount learned from the demonstration was substantially reduced.

Incidental learning, which might be considered to occur in the absence of covert responding, has been demonstrated to depend on (a) self-induced covert sets to learn (170); and (b) responses of selective attention and of silent pronunciation (159). The experimental literature points to the conclusion that when learning occurs, as measured by changes in behavior and in the absence of overt responding, one may be certain that some form of covert responding has occurred.

A theoretical argument in favor of some form of covert responding as a basic condition of learning is that learning occurs in the central nervous system and not in the sense organs and muscles. The stimulus input goes through a mediating central process before an observable act is instigated. This is typified by Hull's model of S-sHr-R, where S is the stimulus input and the R the output. The connection (H for habit) is formed between s and r. A similar view is that the first response to a stimulus is a perceptual one which has stimulus properties that can elicit an overt response. But whatever the nature of s and r may be, they serve as mediators between the S and the R. These mediators may be short-circuited out in the case of reflexes and highly practiced acts. If this theoretical view is adopted, it follows that in learning, some form of covert responding always precedes overt responses, and that the association is due basically to the s-r connection.

This view is supported by the results of a great many experiments on variables that intervene between stimulus reception and overt responding. Examples are Sheffield's "perceptual blue prints" (188), Osgood's response of meaning (166), strategies of associational learning improvised by subjects (24, 25). In an experiment on paired-associate learning using nonsense syllables as pairs, Bugelski (24) questioned his subjects on their learning strategies and discovered that many of them had been quite ingenious inventing mnemonic aids such as "deputize" for "dup-Tax."

This, however, is not the place to review all the experimental evidence on the mediating effects of perceptual, silent verbal, and other forms of covert responding. It is sufficient to note that in learning situations overt responses which are not preceded by a percept, thought, idea, or some other unobservable form of inner activity are exceedingly rare. Thus when comparisons are made between the effects on learning of instructing one group to write the response word, or to vocalize it, or press a button indicating it, and instructing another group to "think" of each correct response, it is most presumptious to assume that the members of the overt group did not first think of each correct response before writing it down.

Such comparisons are actually between covert-overt responding, and covert only. The question arises as to whether overt responding is anything more than giving expression to covert responses. The answer to this question depends on the extent to which connections already exist between thoughts and actions. As E. L. Thorndike long ago (218) pointed out in a paper on ideo-motor action, "An idea has no power to produce an act save the power of physiological connections born in man, or bred in him as the consequence of use, disuse, satisfaction and discomfort."[3] Thus the idea of an act will not function as a stimulus to elicit the act unless a connection has been formed between the idea and the act. For example, however hard one may think of wiggling one's ears, the thought will not produce the wiggle, unless a great deal of practice has occurred. Thus it would appear that overt responding is an essential condition of learning when the to-be-learned response is not in the learner's repertory and not previously connected with a covert response of perception or cognition. This is sometimes called "response learning" as distinct from associate learning (222).

The same idea or thought may become associated with several different forms of overt responses. Expressions may be verbal- oral or written. They may be in the different languages known to an individual. In programmed instruction covert responses (thoughts) may be made manifest either by writing a word in a blank space, by checking a multiple choice alternative, speaking the response word aloud, or pressing a selector button. If the same covert response acquired by association with a stimulus can be overtly expressed in several ways, one would expect to find that the mode of expression has little or no effect on associative learning. The experimental literature tends to confirm this.

In programmed instruction Fry (61) found a slight but consistent advantage of constructive over multiple choice responding during training, but on the criterion test the mean score on multiple choice items was higher than on recall items regardless of the form of training. Coulson and Silberman (38) covaried response mode with size of step and with branching and found only slight differences in mean scores between constructive and multiple choice response modes. Roe (177) found no significant difference between constructed and multiple choice responding in a program of freshmen college mathematics. Hough (101) found no difference between constructive and selective response modes in teaching a course in education to college juniors and seniors. Evans (51) found no difference between these two overt response modes to a linear program designed to teach how to construct short deductive proofs. Evans, Glaser, and Homme (53) compared three modes of responding to a program on symbolic logic – constructive, multiple choice and covert. No significant differences were found on the criterion test, but fewer learning errors were made by multiple-choice responding. Kanner and Sulzer (107) compared overt-vocalized with over-written response in reviews session in learning the phonetic alphabet and found no significant difference in mean scores. Spoken, written and written-spoken responding was comparedon an 87 frame program on electricity by Alter and Silverman (4), with no significant differences in mean scores on a criterion test.

The results of these experiments clearly indicate that when several kinds of overt response have been associated with an idea or thought it makes no difference on the criterion tests which one is called for by the instructions. When these results are coupled with those of a dozen or more other experiments which indicate that, in programmed instruction and in learning from other forms of presentation, instructions to "think" the right responses are sometimes as effective as instructions to express them, it would appear that, except for response learning and other forms to be mentioned later, overt responding is not an essential condition of associative learning.

There are, however, exceptions to this rule. One is that the response mode should be appropriate to the learning task. If for example, the task is learning to spell, either orally or in writing, the response should be either oral or written. But if the task is to recognize misspelled words in print, the optimal mode is probably multiple choice. The same is perhaps true for other forms of discrimination learning. If the task is learning to type, the optimal response mode is obviously that of striking the keys of a typewriter.

2. Advantages of Overt Responding.

Despite the fact that overt responding is not necessary for all learning tasks it has a number of practical advantages for programmers, film producers, and producers of other kinds of instructional materials, as well as some theoretical values for learners.

(a) Advantages for producers. First, records of overt responses have a great value for the experimental production of instructional materials. This advantage is utilized more by producers of programs than by producers of films and textbooks. Most good programs are experimentally developed. Preliminary versions are tried out in classrooms and revised. In order to revise a film or a program, the producer must have available data derived from the try-outs. Such data are provided by the error rate of a program, or by analyses of pre- and post test scores as in the case of films, particularly when test items are constructed to match specific portions of a film. In building a program, the experimenter can manipulate such variables as size of step, number of frames, order of presentation, amounts of prompting amounts and kinds of feedback, forms of overt responding, the number of examples in relation to rules, and so on for other variables. In order to manipulate these variables on which learning presumably depends, the program builder needs a great deal of factual information derived from try-outs which is provided mainly by records of overt responses made during learning as well as on the post tests.

A second advantage accrues mainly to those who wish to construct a branching program, or computer-based program. Unless overt responses are made to each frame and recorded, it is impossible to know what the next step should be, particularly when a computer is employed. It is possible that a branching program could be theoretically constructed without previous data, but it could hardly be operated, especially with a computer, unless there is overt responding.

A third experimental advantage of overt responding arises in connection with the requirements for obtaining continuous learning curves. For example, in the study by Cook and Spitzer on prompting *vs*. confirmation (36), a test trial was inserted after each three prompting trials in order to obtain a learning curve that was comparable to the one obtained by the confirmation method. Furthermore, if an experimenter wishes to measure the time between the appearance of the stimulus and the response, the response must be overt and recorded.

(b) *Advantages to the learner*. These advantages are mainly theoretical but some have been experimentally verified.

First, overt responding gives the student practice in doing what he will eventually be expected or required to do. Learning is conceived as a process by which students pass from an *initial* ability to perform through a stage of *intermediate* behaviors to a stage of *terminal* ability. Gains in abilities from the initial to the terminal stage are measured by the difference in ability to perform on a pretest and on a posttest. The acquisition of knowledge is operationally defined in terms of some kind of performance. Skinner (202) has emphasized the point that the way to impart knowledge is to teach the behaviors from which knowledge is inferred. Covert responses to instructional materials are always *inferred* from observable antecedent conditions and consequent performance, i.e., by comparing what the learner could do in a situation before with what he could do afterwards.

The argument that learning to perform the terminal acts by successive steps or stages is best achieved by overt responding is clearly seen in the acquisitions of skills such as handwriting, typing, using a slide rule, playing a musical instrument, or speaking a foreign language. Also in memorizing a poem the student must practice what he may eventually have to do – recite it. There are other learning tasks, however, for which the need for overt practice is not so apparent. To this point we shall return later.

A second advantage claimed for overt responding during learning is that it provides the experimenter, teacher, or programmer with an opportunity to confirm or reinforce correct responses and to correct wrong ones. If reinforcement is an essential condition of learning, as Skinner and others claim, and if a response cannot be externally reinforced until it is emited, as Holland claims, it follows that overt responding during learning will have an advantage over covert responding. This proposition, however, has been challenged as will be seen later.

A third advantage that may accrue to the learner is that motor acts produce proprioceptive stimuli which, if distinctive, may function to cue correct responses on posttest. In the test situation the visual stimuli presented by the questions or item may elicit incipient muscular responses which produce stimuli that were conditioned to the correct overt responses in the learning situation.

Fourth, while responding overtly the learner has time and an opportunity to make additional covert practice trials. This is most likely to happen when the responses are written.

Fifth, the requirement to respond overtly may have a positive motivating effect. When a student leaves a record of his work which can be inspected and even "graded" he may be more motivated to pay close attention, to read carefully, and to persist longer against the effects of fatigue and boredom than he would if his responses are known only to himself.

A successful method of teaching college freshmen to read and remember is for the instructor to require the student to read a sentence and then close the book and recite in his own words the essential content. When this task has been achieved the instructor requires the reading of a paragraph and a recitation of its contents. This procedure motivates the students to read closely and carefully with intent to learn and to remember.[1]

Despite these numerous theoretical advantages of overt responding, the experimental literature indicates that under some conditions and for some learning tasks overt responses contribute little or nothing to posttest scores over and above that contributed by covert responding.

---

[1] The author is indebted to Dr. Fred Sheffield for this item of information.

3. Disadvantages of Overt Responding.

First, it requires more time than covert responding. In practically all experiments in which learning time has been recorded, the amounts of knowledge gained per unit of time from covert responding have been significantly more than for overt responding. Thus it would appear that for most forms of verbal learning covert responding is the more efficient method. Even in cases where overt responding is the more effective (i.e., results in greater gains), the extra time required offsets this effectiveness when learning is assessed in terms of amount gained per unit of time. There are exceptions to this rule as found in the results of experiments in which learning time was the same for both forms of responding (162, 163).

Another possible disadvantage is that the requirement for overt responding could have, under certain conditions, a demoralizing and discouraging effect. Under other conditions it could have an effect of boredom (153).

4. Factors Related to Response Modes.

Whether overt responding, especially in the second phase of the instructional process, will facilitate, inhibit, or have no effect on learning as measured by posttest scores depends on a variety of factors and conditions some of which can be manipulated experimentally. In this paper they are treated under four main headings, as follows: (a) pre-presentational factors, (b) presentational (c) post-response, and (d) individual differences.

(a) Included in the category of pre-presentational factors are all instructions, directions and information given students before the materials to be learned are presented. Such instructions, when understood and followed faithfully, as they may or may not be, will have an effect on response modes.

(b) Presentational factors include all aspects of how the materials are presented. One important variable is whether the presentation includes one, two, or all three phases of the instructional process. Another is the length of the lesson, and a third is pacing and other timing variables.

(c) The contributions to learning of post-response events depend in part on whether the answers were written, multiple-choice or "thought," and in part on the kinds of covert response that are made to post-response stimuli.

(d) Individual differences which have the most determining effect on the contributions of overt responding to learning are differences in entering behavior and abilities.

II. PRE-PRESENTATIONAL FACTORS

Before presenting the material to be learned the experimenter (or the teacher) usually informs the students concerning (a) the nature of the learning task (b) the response mode or modes to be employed, and (c) and other pertinent information and instructions about the purpose and the procedure of the experiment. These preliminary instructions have various determining effects on optimal response modes.

1. Definition of the Learning Task.

If the task is to acquire a perceptual-motor skill such as handwriting, typing, tying knots, playing a musical instrument, or reading a slide rule, it is certain that overt practice will facilitate the to-be-learned performance. Such tasks also require "perceptual" learning. The performance as it is presented must be carefully observed and understood. Failure to perceive correctly the meaning of the material will interfere with the beneficial effects that may be derived from overt practice.

Learning tasks are determined by educational objectives, specified by the materials to be learned, and illustrated by items in the posttest. One way to clarify the nature of the learning task is either to give a pretest or to give sample items from the posttest. If the task is to acquire a manual skill, a few practice trials may be helpful.

It is important that each educational objective be announced in terms of abilities to be achieved rather than merely in terms of subject-matter to be learned. This point has been stressed by Mager (143), Skinner (202), Lumsdaine (128), Taber, Glaser and Schaefer (215), by Gagne (66), and others. Learning, as noted above, is viewed as a process by which the learner moves from an imperfect initial performance (called "entering behavior") through intermediate learning activities to a more perfect terminal performance. The optimal response mode during the learning stage depends not only on the nature of the terminal performance but also on each individual learner's entering behavior.

Gagne (66) has proposed a taxonomy of behavioral objectives and a related taxonomy of types of learning (Gagne (65). The attainment of each of his seven categories of objectives requires a different type of learning and each type of learning depends on a different set of conditions. Although Gagne does not include any particular response mode as one of the conditions of learning for any one of his seven types, yet for some of them it is apparent that an overt mode is tacitly assumed to be one of the conditions.

His first behavioral category is response differentiation, and the corresponding form of learning is response learning. The task is to add new responses to the learner's inventory and to discriminate these new responses from each other and from older ones. Examples are: learning to handwrite by copying a model, learning to speak a foreign language by imitating speech sounds, and learning to produce a sound of a given pitch by imitating a sound made by a piano key or a tuning fork. This form of learning requires the establishment of point to point correspondences between stimuli and responses. It results in what Lane (123) has called formal repertoires of responses which are distinguished from thematic repertoires of response in which the responses are not controlled by the stimuli that they represent. But formal repertoires are basic for the attainment of thematic ones.

The optimal mode for response learning is overt, as will be discussed more fully later in this paper. Unless the response is uttered, written, or otherwise expressed, it would be impossible for an experimenter or a teacher to reinforce right responses or nonreinforce wrong ones.

Gagne's second behavioral category is association and the related type of learning is associative, commonly known as S-R learning. In this form of learning the response is not a copy of the stimulus, as it is in the first form indicated above, but functions to name the stimulus or otherwise code it. One of the basic conditions on which associative learning depends is that the to-be-learned response is already in the learner's repertory. When this and other conditions on which associative learning depends are fulfilled, covert responding may be just as effective as overt and perhaps more efficient from the standpoint of the time required to accomplish the learning task.

For the purposes of this paper, it is neither necessary nor profitable to run through the remainder of Gagne's categories and speculate on what the optimal response modes for each might be. This is a promising area for future research. It is sufficient to note that the optimal mode for some of the behavioral categories and their corresponding forms of learning tasks may be overt, and for others it may be covert. It is important to note, however, that his types of learning are arranged in a hierarchy so that one of the conditions of learning at each level, except the first, depends in part on having mastered the tasks set by the lower levels. If, for example, the response learning that is required at any level has not been previously acquired, the optimal mode at this level may be covert plus overt; whereas if the needed response learning has been acquired, the optimal mode may be only covert.

2. Response Modes in Incidental Learning.

When the learning task is left undefined, or when an orienting task other than instruction to learn is given, the effects on learning are reported in a number of experiments on incidental learning. These experiments indicate that in the absence of any externally induced set to learn a certain amount of learning of verbal material does occur (168). The amounts learned incidentally depend on a number of factors such as the meaningfulness of the materials and parts of it to which attention is given.

Incidental learning is attributed by Postman and Senders (170) to self-induced covert sets which transfer by generalization from previously acquired learning sets. Such covert sets vary in strength depending on the degree of transfer. The theory of incidental learning as advanced by Postman (168) is that learning sets, whether induced externally or self-induced, are independent variables which control the kinds, frequencies and distributions of responses (overt or covert) that are made to the materials. Such responses have been called "representational" because they are indentical in linguistic form to the presented stimuli (46). The frequency of responses controlled by self-induced sets is less, and their distribution is different, from those controlled by externally induced sets. This accounts for the superiority of intentional over incidental learning. Representational responses may be either covert only or covert plus overt. Very few experiments have been reported in which subjects were forced to respond overtly under conditions of incidental learning. The earlier ones (21, 104, 169) indicate that even when responses are overt, intentional learning is better than incidental. In a more recent experiment, Mechanic (157) found that when incidental learners can be induced to make high frequency pronunciation responses, they learn about as well as under conditions of intentional learning; but with a low frequency of pronunciation responses, they learn only about half as well as under conditions of intentional learning.

In another experiment (Mechanic and D'Andrea (158), five different conditions of articulation were covaried with incidental and intentional learning. They were (1) to spell each word in a list of trigrams silently, (2) aloud (3) pronounce each one silently (4) aloud, and (5) merely to look at each item during exposure. Each list of trigrams was composed of nonsense syllables and high and low frequency words. The orienting task for all subjects was to guess which of the three letters in each of the trigrams had been chosen to be "correct" and to call it out when a trigram was exposed. Half the Ss were told that they would be expected to recall the trigrams (intentional learners) and half had no such instructions (incidental learners). Each item was exposed for three seconds during which the articulation instructions were performed, and the guess was enunciated during a one second inter-item interval. Each list of 20 items was presented twice, after which all Ss were given a five minute test of free recall. Score was the number of items that could be correctly recalled.

Results were: (1) Overt responding was definitely better than covert for intentional spelling, but not for incidental, and better for pronunciation in both groups. (2) Intentional was better than incidental for all treatments except overt pronunciation, where the mean number of items recalled was about the same for each group. (3) There was no interaction effect with the meaningfulness of the trigrams. All groups recalled high frequency words better than low frequency ones and all words were recalled much better than nonsense syllables.

Overt spelling was better for intentional than for incidental learning because incidental subjects who were instructed to spell silently were less likely to do so than intentional subjects. The results, however, tend to confirm the hypothesis that, in actual practice, overt or covert responding is a more potent factor in learning than "set" to learn.

The fact that there was no significant interaction between meaningfulness of items and response modes or learning "set" led to a further experiment by Mechanic, (159) to test a hypothesis advanced by Postman (168) that incidental learners respond to fewer stimuli than intentional learners. It has been observed that incidental learners do better on more meaningful items than on nonsense items, while this is less true for intentional learners. Hence the difference between incidental and intentional learning tends to decrease as the meaningfulness of the materials increase. However, Mechanic (159) argues that when subjects are set to learn they make more pronouncing responses than otherwise. It is these responses that account for the difference between incidental and intentional learning of materials that vary in meaningfulness. Without a set to learn there is a greater selectivity by subjects in regard to their pronouncing responses. Instruction to learn tends to interfere with this selectivity.

To test this hypothesis two experiments were performed. In both, the subjects were presented with a mixed list of 12 high and 12 low frequency trigrams. The score was computed by dividing the difference between highs and lows that were recalled by the total recalled, (H-L/H + L x 100). In one experiment, both intentional and incidental learners were required to pronounce each item as it appeared on a memory drum. Each list was repeated eight times in rapid succession after which a five-minute free recall test was given. The difference between the mean scores of the incidentals and intentionals was not significant. The second experiment was like the first in all respects except that no pronunciations were required, and the orienting task was to guess whether or not each item was in a sealed envelope -- as though it was an experiment in extra-sensory perception. The mean score of the incidentals (H/L total) was significantly higher than that for the intentionals, meaning that the incidentals paid much closer attention to the high frequency words than to the low frequency ones. So the phenomenon of selectivity for high frequency words disappears when overt responding is required.

3. Response Mode Directions.

The pre-presentation variable that has the most direct effect on response mode is the instructions given to subjects as to how they should respond. In a typical experiment one group of subjects is instructed to respond to frames in a program by writing correct answers in blanks, or by checking an alternative on multiple-choice items, or by pressing a selector button, or by articulating a verbal response, or by engaging in some other specified form of overt responding. Another group may be instructed to "think" the correct answers and not respond overtly, and still another group may be instructed only to read the materials. The results of such experiments may depend, in no small way, on how these instructions are actually understood and obeyed by all subjects. The experimenter cannot be absolutely certain that members of the "think" group did actually think the answers, or that the "read" group did read carefully and thoughtfully, or even that everyone in an "overt" group did take the time to write all the answers, or remember the instructions not to skip any.

Failure to follow directions on the part of some subjects can foul up results of an experiment. Kimble and Wulff (112, Ch. 16) discovered this in one of their experiments on the role of active participation in learning to read the "C" of "D" scale of a slide rule. Replicas of the scale were printed in booklets. The task was to draw a line across the scale at a point that corresponds to a given number. Each response was guided or prompted by two parallel lines drawn some distance apart across each scale between which the correct answer could be found.

One group of subjects got booklets with all answers correctly marked and was told to study them. They had zero percent of participation examples. Other groups got 25%, 50%, and 100% of participation examples to which they were instructed to respond overtly by marking the correct answers. The hypothesis was that the greater the percentage of participation, the higher would be the mean score on the posttest. But the results showed the highest score for the zero participation group, the next highest for the 75% group, with the 25%, 50%, and 100% groups as poor thirds.

In commenting on this result the authors say: "It was evident from the individual participation booklets that many subjects did not participate when instructed to do so. It is also apparent from the booklets that many of the nonparticipation subjects did participate actively without instructions to do so...these failures to follow instructions make it difficult to determine treatment test means that accurately reflect the true means for the treatments investigated." (p. 238)

Fortunately, these investigators took the trouble to examine the records to find out how well instructions had been followed. Failure to do this on the part of experimenters who have compared overt and covert modes could cast a shadow of coubt across their results.

Furthermore, the directed response mode may not be pure in the sense of being uncontaminated by other modes. For example, members of "think" and reading groups may not only think but also emit a variety of incipient motor responses that are not recorded. Among such responses are eye movements, muscle potentials, EEGs, GSRs, and neuroelectrical brain reactions. Instruction to do nothing but read or think does not prevent many implicit and incipient motor responses from occurring (152). Some psychologists hold that there is no such thing as a pure covert response that is wholly devoid of some externally measurable and concurrent manifestation. Likewise all observable overt responses are preceded or accompanied by thought, except reflexes and highly practiced acts. In learning situations, overt responders surely think before they act, or while they are acting. Thus it appears that in experiments where posttest scores of groups who were instructed to respond overtly are compared to scores of those instructed to "think" or "read," the results are only crude manifestations of the gross effects of pre-presentation instructions. Aside from whatever heuristic and practical value such experiments may have, they fall considerably short of providing a clear understanding of the relation between how learners respond to a presentation and what they learn from it.

Despite the fact that some experimental subjects may misunderstand preliminary instructions, forget them, or fail to follow them faithfully, and despite the fact that all response modes may be mixtures of overt and covert elements, pre-presentation instructions and information are known to have determining effect on learning, however gross they may be.

We know that in learning from films, prefilm instructions do exert a considerable influence on what is learned and how well it is learned (150). The effect is produced partly by directing attention to relevant aspects of the materials and partly by increasing the motivation to learn.

4. Pre-presentation Motivators.

In a previous paper (150) where the experimental work on the effects of directed attantion on learning from films was reviewed, it was found that prefilm motivators have a greater effect on learning than built-in motivators. This effect was attributed to the fact that prefilm instructions defined the learning task more specifically. Motivation is most effective when it is relevant to the learning task.

Explicit instructions to respond overtly could have a greater effect on motivation than instruction to "think" the answers because the subject leaves behind a record that may be inspected by the experimenter or the teacher on which his performance is graded or otherwise evaluated. If a subject only thinks the wrong response no one else is the wiser, but if he puts it on record he may be corrected or perhaps criticized. When subjects are instructed to respond verbally by calling out their answers aloud in the presence of others, as well as in the presence of the experimenter or teacher, this social situation might act to increase motivation to pay closer attention and to try harder to remember the correct answers.

In such a situation motivation may be derived from a desire to excell or to show-off in the presence of one's peers.

Explicit instructions to respond overtly could have, under some conditions, a deleterious effect on motivation. In the early stages of learning when subjects are likely to make a good many errors, the requirement to leave a record of them could have a demoralizing effect. On the other hand, if the questions or frames are too easy, subjects may become tired of writing them down and bored with the whole task.

The experimental literature on the possible motivational effects of explicit instructions is not very extensive and is far from conclusive. The classic experiments on the effects of instruction to respond orally in group situations are those on learning the phonetic alphabet reported by Hovland, Lumsdaine and Sheffield (100) and by Lumsdaine and Gladstone (134). A letter and its phonetic equivalent were presented one at a time on a screen. After each six to eight pairs were presented a review frame was introduced. For one group of subjects - called the "active review group" - each letter was shown one at a time and the members of the audience were instructed to call out the correct phonetic word. The "passive review group" was shown both the letters and their words with the experimenter pronouncing the phonetic word just as he did in the learning session. In the testing situation the letters were presented on the screen one at a time and the subjects were required to write the correct phonetic word on a worksheet. In both experiments it was found that the active review groups scored significantly higher on the posttest than the passive review groups. Furthermore, it was noted that the subjects in the active group who were the least motivated and slowest in learning the more difficult items profited the most by active review. However, in the later experiment by Lumsdaine and Gladstone (134), using the same film and the same experimental procedure, the low ability active group gained less over the low ability passive group than the high ability active gained over the high ability passive.

The possible motivating effect of the instructions to call out each response by the whole group might be attributed, in some part at least, to a competitive factor, but this is obscured and confounded with other possible effects. One is that the subjects who did not know the correct response waited until they heard it from someone who spoke up promptly, and thereby gave a prompted reply. When the experimenter announced the correct answer, after a few seconds, this served as a confirmation.

If a subject is required to respond (or recite) orally in front of a class, instead of in concert with others, this requirement could have an anxiety-producing effect which would interfere with learning. As pointed out by Taylor and Spence (216), a small amount of anxiety may have a positive effect on performance, while high levels may interfere with learning.

The motivating effect of an announcement that a posttest would follow the presentation of a film was investigated by Michael and Maccoby (163) and by Maccoby, Michael and Levine (139). They found that such an announcement had no effect on posttest scores for a film that was intrinsically interesting. But for a film that lacked intrinsic interest there was a positive and significant effect. This result is in accord with the findings from other experiments on the effects of extrinsic motivators (150).

In an experiment on learning the names of symbols that appear on maps, Kendler, Kendler and Cook (110) covaried participation instructions with sex, and I.Q. Four participation conditions were tried: (1) the _S_s were instructed to write the correct word when symbol was displayed; (2) _S_s were told to think of it but not write it; (3) no instructions were given but a maximum opportunity and time for responding were allowed, and (4) no instructions were given, but the time allowed in condition (3) was filled up with noninformative verbiage which was intended to interfere with or prevent any covert respondings. The results showed that the two "instructed" conditions were better than the two noninstructed, but there were no significant difference between instructions to write or just to think. But the noninstructed maximum opportunity condition was definitely better than the minimum opportunity condition. The authors suggested that the superiority of instructions versus no instructions could be attributed, in some part at least, to a motivating effect.

The motivating effect of instructions is also indicated in the results of an experiment on problem-solving by Gagne and Smith (70). The problem was to transfer a stack of graduated discs from one place to another so that the order of size in the new stack would be the same as in the original. One group of students was instructed to tell the experimenter why each move was made and also to try to think of a rule that could be told to some other person. A second group was instructed only to announce a reason for each move. A third group was told only to try to think of a rule, and a fourth group was given no specific instruction beyond following the rules of the game. There were four practice sessions in which the problems were to transfer, 2, 3, 4, and 5 discs. The criterion test was a 6-disc problem. The two groups who were instructed to verbalize a reason for each move during training had significantly higher mean test scores than the other two groups, both in the numbers who got a perfect solution and in time taken on the criterion test. The authors concluded that the instructions to verbalize had a motivating effect of forcing the Ss to think up new reasons for their moves.

The motivating effect on attention and effort to learn resulting from instructions to respond overtly and from provisions for doing so has not been sufficiently explored to justify a firm conclusion. It is apparent, however, particularly from the studies of Michael and Maccoby (163) and from those of Maccoby, Michael, and Levine (139), that when the level of motivation from other sources is relatively high, instructions to practice add very little to the existing level of motivation. One source of motivation is a clear knowledge and understanding of the criterion test. If a subject knows precisely what he is expected to learn, he will have in mind what John Dewey called "an end in view." For example, in learning a skilled performance from a demonstration, the task to be learned is clearly defined. Likewise in learning from programmed materials, the tasks are presented step-by-step.

Another pre-presentation variable that may have an effect on response modes is an announcement concerning time limits. If students are instructed to respond overtly and also to work against a time limit, some *of them may be tempted to ignore or forget the response instruction.* The effects of time factors on response modes will be discussed later in this paper.

In summary, pre-presentation variables such as (a) clarifying and illustrating the nature of the behavioral objectives and the learning task, (b) response mode instructions (c) the announcement of a posttest and other instructions and explanations, could have varying effects on optimal response modes. The effects will depend on how well these instructions are understood, remembered, and obeyed.

III. PRESENTATIONAL FACTORS

The contributions that overt responding may make to learning depends to a large degree on how the materials are presented. When presented as continuous discourse, overt responding such as reading aloud contributes little or nothing to learning except perhaps under very unusual conditions. But when the materials are programmed or semi-programmed and presented in two or all three of the instructional phases, overt responding in the second and third phases may contribute substantially to learning.

In programmed instruction the three major phases are reading, answering, and confirming. Instructional films and ETV programs may be programmed by dividing the presentation into predetermined periods of observation and reading, and intervals for questions and answers, and presentation of correct responses. In addition to phasing there are at least two other presentational factors that may have a determining effect on the contributions of overt responding. They are the length of the lesson, the rate at which the parts of the materials are presented, and other timing variables.

1. Relation of Response Modes to Instructional Phases.

In each of the three phases the inevitable responses of attention, perception and cognition may or may not be supplemented with some form of overt responding. The contributions which overt responding may make to correct perception, attention and understanding in the first phase has not been investigated enough to warrant a review of the evidence at this time. It is at the second phase where most of the experimental work on response modes has been done. At the third phase very little has been done on whether or not learning is facilitated by requiring students to pronounce or write the correct responses to frames as they are presented. This problem will be considered in the next section of this paper.

The research on response modes in the second or answering phase has been done almost entirely on linear programs. All branching or "scrambled book" programs, and all computer-based programs, as well as some that are prepared for use on teaching machines, require overt responding in the second phase because it is an essential condition for the administration of such programs. It may be noted, however, that a student could "think" his way through a "scramble book" program and do nothing overtly but turn the pages. But no experiment, to the knowledge of this writer, has been reported on a comparison between amount learned from thinking through vs. working through such a program.

All tutorial types of instruction that are based on interactions between student and teacher, student and machine, or student and computer require overt responses from both parties. The accomplishment of many learning tasks is undoubtedly facilitated by active give and take between student and teacher regardless of whether the teacher is a person or a machine (124). The values of adaptive systems of instruction will be considered later in the section of this paper on individual differences.

The contributions which overt responding in the second phase of the instructional process may make to learning have been investigated in connection with: (a) instructional and training films, (b) programmed ETV instruction, and (c) printed linear programs.

(a) Overt responding to instructional and training films. In the earlier experiments on learning from films and other AV presentations, comparisons were made between amounts learned from viewing the film as a whole and amounts learned when the film was stopped at intervals for periods of review, practice or "active participation." In a typical experiment one group of subjects sees the unprogrammed version of a film and a comparable group sees the programmed or semi-programmed version. Learning from the unprogrammed versions is limited to learning from only the first of the three instructional phases; while learning from the programmed versions is usually from all three of the phases of observation, answering, and confirmation. The results of these experiments indicate that groups of students who see the programmed versions tend to score significantly higher on posttests or make relative greater gains from pretests to posttest. The programmed versions, however, usually require more time. This may be equalized by showing the unprogrammed versions twice, in which case differences in mean posttest scores are often reduced. The experiments referred to here are reviewed by Allen (1), Lumsdaine (130), and Lumsdaine and May (135).

The question arises as to whether the superiority of the programmed version of instructional and training films is due to the fact that in most experiments the responses in the second phase were overt, or to some other factor or factors such as opportunities for additional covert practice or rehearsal trials. The answer to this question depends partly on the nature of the learning task. If the task is only to acquire new information, there is some evidence that covert responding with knowledge of correct results (KCR) is as effective as overt responding with KCR. This evidence comes from an experiment by Michael and Maccoby (163) who used a film called "Pattern for Survival" on civilian defense against atomic bombing. Running time was 13 min. and 44 sec. It was stopped at three intervals, and questions were asked orally by the experimenter on some of the points covered in the preceding section. The overt-responding groups wrote their answers on practice question sheets, and the covert groups were instructed just to "think" the answers. The posttest consisted of 30 orally administered questions. Half of the questions had been practiced during the participation periods and half had not been. A control group saw the films without interruptions for practice. Another control group took the test without having seen the film. The mean score of this group was about one half as great as that of the groups who saw the film only. The participation groups scored significantly higher than the "film only" groups on the practiced test items but not on the non-practiced items. The mean scores of the groups who practiced overtly were about the same as those who practiced overtly. This was true for both high and low I.Q. students. But the groups who received KCRs during practice (the third phase of the instructional process) scored significantly higher than the groups from whom KCR were withheld.

If the learning task is the acquisition of a new skill rather than new knowledge, the experimental evidence indicates that overt responding does make a substantial contribution to learning over and above that made by merely "thinking" the correct responses. In all such tasks the student is required to reproduce the stimulus materials presented during the reading or observation stage. This is illustrated by the experiments on learning the phonetic alphabet where it was found both by Howland, Lumsdaine and Sheffield (100) and by Lumsdaine and Gladstone (134) that overt active review sessions produced better learning than "passive" review sessions. The passive review groups, however, were not instructed to "think" the associations. In a later experiment on learning the phonetic alphabet, Kanner and Sulzer (107) found that when the "coverts" were so instructed, the superiority of overt oral responding was substantially reduced. In other experiments where the responses to be learned are nonverbal or where new verbal sounds are to be learned by imitation, overt practice appears to be an essential condition. This point will be discussed later in connection with response learning.

(b) Overt responding to ETV presentations. Heimer (92) compared three different methods of presenting a programmed course in college algebra with conventional teaching of the same content. The three programs were: (a) book form and self-paced (b) teaching machine, self-paced, and (c) group presentation on filmstrips, externally paced. The conventional teaching paralleled the unit sub-divisions of the programmed materials and occupied about the same amount of time per unit. Tests were given at the end of each of the 15 units plus a final examination. All responses to the programs were overt. The percentages of correct answers on 14 of the 15 quizzes and on the final exam were greater for the programmed groups combined than for the conventionally taught groups. The conventionally taught groups had a mean score of over 75% on only three of the unit tests, while the programmed groups reached or exceeded this average on 13 of the unit tests. The variances of the test scores of the conventionally taught group were considerably greater than the variances of the program-taught groups. Again the question arises as to whether or not the superiority of the program-taught groups could be attributed to the fact that their responses to the frames of the program were overt.

Some light was thrown on this question by a series of experiments by Gropper and Lumsdaine (86, 87, 88) where televised presentations were constructed in ways to permit overt responding and compared with conventional teaching of the same materials. In the televised structured versions the teacher would pause before completing a sentence, or before an expected response to allow time for the students to "think" or to write a word or phrase that would complete the sentence while a question mark flashed on the screen. After a few seconds the teacher would supply the correct response. One group of subjects saw the conventional version in one room while another saw the programmed versions in another room. A posttest was given immediately and again after a delay from 10 days to two weeks.

Results were obtained on two different lessons -- one on heat and one on nuclear reactions. The subjects were 7th and 8th grade students. On the posttest for the lesson on heat students with high I. Q.'s had a higher mean score on the programmed presentation both on the immediate and delayed tests. For students with low I. Q.'s mean scores on immediate test group were not different for the two presentations but on the delayed test the conventional groups had significantly higher scores. For the lesson on nuclear reaction the groups who had the programmed version had higher mean scores than the groups who saw the conventional version in five out of six comparisons.

It may be noted that in this experiment the responses made to the conventional presentation were implicit-covert[1] throughout, while those made to the programmed versions were a mixture of implicit-covert and explicit-overt or covert, depending on whether the subject chose to record or only to "think" the answers. Thus it would appear that making provisions and giving instruction for explicit responding (either overt or covert) did add an increment to learning over that obtained by implicit-covert responding alone.

In another experiment using a lesson on "Newton's Laws of Motion," no significant differences, either on the immediate or delayed posttests, were found between mean scores of groups who saw the conventional and programmed versions. This was interpreted to mean that encouragement to respond to the pauses without assurances that the responses would be correct contributed nothing to learning. The effectiveness of "active" responding would appear to depend on the response being a correct one.

To test this hypothesis another experiment was done using a lesson on "How Movies Work." The programmed version was sequenced in small steps and were so well prompted as to almost guarantee that correct responses would be made. Now the results were different. The students who had the programmed lesson had a higher mean score on the posttest than those who saw the conventional version.

---

[1] The terms explicit and implicit refer to pre-presentional instructions. Implicit-covert means that the students were not instructed to think the answers.

A comparison between explicit-overt and explicit-covert responding was made using a lesson on body chemistry. One group of subjects was instructed to write the response on a work sheet (completing the sentence left incomplete on the screen, while a control group was instructed to read the sentence silently while the instructor read it twice. No significant difference appeared between mean scores of these two groups either on the immediate or delayed posttest.

The results of these experiments suggest that (a) instructed-overt responding is no more effective than instructed-covert responding when provisions are made for them, and (b) the contribution to learning made by these two response modes is due in some part to the elicitation of the correct responses. This would seem to imply that if a subject read or observed carefully so as to make the correct perceptual responses, the implicit-covert mode would be about as effective as the explicit modes.

(c) Overt responding to linear printed programs. Here again comparisons have been made between amounts learned from reading continuous discourse and from programs based on the same materials. Evans, Glaser and Homme (53) compared the amounts learned from reading ten pages of a standard textbook on statistics with amounts learned from a linear program based on these pages. The mean test score of the program instructed group was slightly but not significantly higher than that of the reading group. But the number of cases in each group was 16 and 17. There was significantly less variance in the individual scores of the program group than in those of the reading group. The experiment was repeated on a linear program using as subject matter the fundamentals of music. This time the mean posttest score of the program group was significantly higher than that of the reading group. And again, the program group had a lower variance among scores than the reading group.

In both experiments the response modes to the program frames were of the overt construction type. It could be true that the contribution which the overt responding requirement made to learning was to motivate a closer and more careful reading of the materials during the first phase of the instructional process. There is further evidence (shown later) that the better the materials are learned in the first phase, the less will be the contributions of overt responding in the second.

At least thirty-nine experiments have been reported in which the effects on posttest scores of overt and covert response modes in the second phase have been compared. Of these 21 report no significant differences on posttest scores or on gains in knowledge between groups who were instructed to respond overtly and comparable groups instructed with to "think" the correct answers or merely to read frames with the correct words filled in. There are, however, at least eleven experiments in which an overt mode has been found to be superior to covert modes, and three or four are reported in which overt responding tends to interfere with learning. As Lumsdaine and May (135) have noted, tallying the results of these experiments does not tell us anything about the conditions under which overt responding contributes to learning, which is what we want to know.

Holland (98), who is an advocate of overt constructive responding, has checked some of the experiments which resulted in no significant differences between covert and overt responding groups against three conditions which he asserts should be met for significant differences to appear. These are: (a) "programmed material must be designed so that the subject can answer correctly; (b) ..... so that he can answer correctly only after engaging in the appropriate mediating behavior; and (c) the program must be long enough for subjects in the covert condition to become careless since, under controlled conditions, they may respond consistently for awhile." (p. 93) He found that a number of the experiments in which there were no significant differences were deficient in one or more of these three criteria.

When a program is designed so that all subjects can respond correctly to most of the frames the error rate is low. The programs used in some of the experiments which resulted in no significant differences between overt and covert responding had low error rates (53, 80, 208); Goldbeck and Campbell (53) found that overt responding to an easy version of a program which had a low error rate actually appeared to interfere with learning. It was on the version with the highest error rate that overt responding was significantly better than covert responding.

Holland (98) notes that there is more than one way to keep the error rate low. One is by selecting subjects to whom the lesson is familiar, another is by over-prompting, and still another by using frames that need not be read at all in order to answer correctly. His second rule demands that the student be able to answer correctly only after having read the frame carefully or otherwise having engaged in what he called "the appropriate mediating or precursory behavior."

In order to test whether or not a lesson is programmed so that correct answers depend on a careful reading and understanding of the critical content of each frame he invented a black-out technique. This technique consists in blacking out all the words (or other stimuli) in a program which can be obliterated without reducing the error rate (Kemp and Holland (109). This technique was applied to the programs used in eight experiments in which no significant posttest differences were found between groups who were instructed to respond overtly or covertly, and to four experiments in which overt responding was the better mode (109). The per cents of the material that could be blacked out without affecting the error rate on the eight programs that yielded no significant differences ranged from 31% to 75%, while the amount that could be eliminated from the four programs to which overt responding was superior ranged from about 11% to 25%. The authors conclude: "If the response is unrelated to the critical content on which the S is later to be posttested, it makes no difference whether, much less how, he responds, but if S is tested on things which have served as a contingency for correct responses, overt responding is important" (p. 113).

This technique was applied to the programs used in about one half of the experiments from which no significant differences on posttest results were found between overt and covert responding groups. Holland (98) noted that among the other experiments there are some that appear to meet his three main criteria. Thus it would appear that the contribution which overt responding makes to learning, over and above that contributed by covert responding, depends in part, but not wholly, on making the correct answers contingent upon appropriate precursory behavior. Again the evidence indicates that the better the materials are learned during the reading and observing phase, the less is the importance of overt responding in the answering phase.

Further research may very well reveal that even when Holland's three criteria are met overt responding may not be any more effective than covert responding and usually not as an efficient method of learning.

As indicated in the preceding section of this paper, the optimal mode of responding may be determined, to some extent at least, by the behavioral objectives and the forms of learning that are involved. For response learning the overt mode is clearly indicated; but for association learning it is not.

There is the possibility that optimal mode of answering responses may be dependent on the response mode employed during the reading phase. The frames in a program may be read silently, as they usually are, or aloud. In paired-associate learning the stimulus terms may be read silently, uttered or even spelled. Answering response may be construction, multiple-choice, written or spoken, verbal or manual.

The standard procedure employed with programmed textbooks requires silent reading of the materials and either constructed or multiple-choice responses to the frames. But these two combinations by no means exhaust all of the possible ways in which response modes at both phases can be varied. Reference was made earlier to an experiment by Mechanic and D'Andrea (158) in which the orienting task was to learn which of three letters of a trigram had been chosen to be correct. The posttest was from recall of the items in each list. The group that was instructed to pronounce each word aloud, and the one instructed to spell it, did better on the posttest than the groups instructed to pronounce it or to spell silently. Mechanic also found that one reason why intentional learning is better than incidental learning is that when subjects are set to learn they make more pronouncing silent responses to the stimulus materials than they do when reading it with no intent to learn.

The relation between reading response modes and answering response modes in programmed instruction has not been thoroughly investigated. Stevens and Sherman (208) report an experiment in which a positive correlation was found between eye fixations in silent reading of the frames, and the number of words left blank to be filled in. This result suggests that the optimal mode of answering responses may, under some conditions, be dependent on the mode of responding to the presentation, and on the media through which the material is presented.

2. Relation of Response Mode to the Length and Structure of a Presentation.

Holland's third condition favoring overt responding is that the program be long enough so that subjects in the covert condition will become careless and skim over the critical materials. He notes that in at least five of the experiments in which no significant differences were found used relatively short programs -- less than 100 frames. But the fatigue or pall effect of working through a program depends not on the amount of material that is programmed but on the length of each lesson. It could be argued that inasmuch as overt responding requires more physical effort than thinking or reading the tiring effect would be greater. Whether the motivation supplied by the requirement of overt responding will be great enough to overcome this effect is an unanswered question. The negative effects of the length of a lesson on learning could be due partly to the generation of reactive inhibition as well as decrease in motivation (184).

The aspect of length which has received the greatest amount of experimental attention is the so-called "size of step" problem. As Lumsdaine (136) has pointed out, size of step is an ambiguous phrase. The sense in which it is here used is the number of words or sentences to be read in the first phase of the instructional process before the second or answering phase is introduced. In an AV presentation, it is the length of a segment of film that is presented before the film is stopped for review, practice or test questions. The response mode in the reading or observational phase is usually, but not always, covert.

The problem is to determine the optimal amount of content that students can comprehend and hold in mind before being required to answer questions about it. The greater the length of the reading or observational phases, the greater the number of questions and answers that may be required in the second phase. This relationship is contingent upon the manner in which the materials to be read or observed are structured. The problem is not one of length alone, but of length and structure combined. Another relation between these two phases is that the longer the first phase, the fewer prompts are provided for correct responding in second phase. One would, therefore, expect that the longer the first phase, the greater will be the errors in the second. Thus it can be said that the greater the amount of reading or observing to be done prior to the second phase, the less detailed are the materials programmed.

Amounts of material to be read or observed in programs may vary from a phrase to a paragraph. Programs of the scrambled-book type may present several paragraphs before an overt response is called for. In learning to perform a manual operation from observing a filmed or live demonstration students could be instructed to imitate each movement as it is demonstrated, as was done by Roshal (178) and by Jaspen (103). This is a minimal amount of presentation or size of step that precedes overt responding. Larger segments were introduced and varied in size in a series of experiments by Sheffield, Maccoby (189, 190) who investigated the relation between amount of demonstration and practice in learning to assemble the parts of an automobile ignition distributor and the parts of an airplane waste-gate motor. In two additional experiments the task was to construct an equalateral pentagon. In each of these tasks it was possible to count the total number of moves required. The distributor required 30, the waste-gate motor 64, and the pentagon nine. The number of moves demonstrated prior to a practice period was determined in some instances by amounts of information that could be held in mind by 75% of a sample of subjects. This was called "demonstration-assimilation span."

In some instances the D-A span corresponded to "natural units," defined as segments or units which have a common set of contextual cues that enable the learner to integrate them readily into a super ordinate unit of the total task. After a demonstration had been segmented either into D-A spans or natural units, it was possible to vary the number of such units that were demonstrated prior to practice periods. The film on the assembly of an ignition distributor was divided into four segments of 9, 4, 7, and 10 correct assembly responses. For one group of subjects, the film was shown in four segments with a practice period after each segment; for another group the first two segments were shown before practice, and the second two were shown and practiced. For a third group the whole film was shown before practice. Each group saw the film and practiced three times. A fourth group practiced after each segment on the first showing, after the first and second segments on the second showing, and after the whole film on the third. This was called the transition group.

After the third showing a criterion test was given consisting of making a correct assembly of the 30 parts. Performance rate was measured by dividing the number of correct moves by total assembly time.

The learning curves of the four groups show sharp drops in assembly errors from the first to the third practice periods. For the "whole" group the drop was from a mean of thirteen to a mean of two errors; for

2/

the larger segment group the mean drop was from about seven to one error, and for the other two groups it was from a mean of five down to a mean of one. Similar drops were posted for mean assembly times. The performance rates of all groups went up substantially. The performance of the D-A group and the transition group was superior on each of the three practice trials and on the criterion test. The "whole" group had mean performance rate of about 2/5ths less than that of the D-A span group. The low performance rate of the "whole" group was due more to the fact that it took more time to complete the assembly than other groups, than to the fact that it made more errors. The error discrepancy was very small. This suggests that its performances both on the practice trials and on the criterion test were more deliberate or "reasoned" than those of other groups who performed as though the task had been well practiced. It would appear that the covert mediating responses functioned faster to elicit the correct responses.

Of particular interest is the performance of the transition group. For the first showing it performed about the same as the D-A span group as would be expected; on the second showing it practiced the same as the double-segment group but performed significantly better; and on the third showing it was parallel to the "whole" group but performed a great deal better. On the criterion test the mean performance score of this group was only slightly less than that of the D-A span group. This suggests that as learning progresses from earlier to later stages the length of the time devoted to learning by observation or by reading or both may be increased without loss in performance rate.

In the experiment under consideration the maximum time that could be spent on the observation phase was 18 minutes (the length of the film) comprising a total of 30 responses demonstrated in an orderly sequence. During the first showing the transition group practiced after having seen 9, 4, 7, and 10 responses demonstrated; during the second showing it practiced after having seen 13 and 17 responses demonstrated; and after the third showing it had seen 30 correct responses demonstrated. In spite of the increase in the length of the phase of implicit-covert responding, the transition group kept pace with the small segment group.

Because of the possibility of wide individual differences in the D-A span, it is possible for each subject to set his own span by being permitted to stop the film for practice at any time. This is called the self-fixing condition, in which each learner decides for himself the size of step or amount of material that he can hold in his head before consolidating it by test trials or practice. This was one of the methods used in a second experiment by Margolius and Sheffield (146), and in an experiment by Weiss, Maccoby, and Sheffield (229) in which the task was one of geometrical construction. In the second experiment by Margolius and Sheffield, self-fixing was as successful as the transition method, even though the subjects did in fact stop the film for practice fewer and fewer times as the number of showings increased.

In the experiment by Weiss, Maccoby and Sheffield (229), where the task was to construct an equilateral pentagon inside of a circle, nine successive steps, which were divided into five D-A segments and into four natural units, were required. In one experiment four groups were compared: one group which practiced after each natural unit, a transition group, a self-fixing group, and a "whole film" group. Each group received three showings, and a criterion test. For a sample of junior college students the "whole" method was definitely inferior to the other three treatments which were about equally effective on the three learning trials. But for a sample of liberal arts and graduate students the self-fixing method was superior to the other three both on the learning trials and on the criterion tests. The subjects who were permitted to choose the number of times they would stop the film for practice chose, on the average, three during the first showings, two on the second and one on the third, which compares with 4-2-1 imposed on the transition group. The self-fixing method has a direct bearing on the question of size of step in programmed instruction.

The results of the foregoing experiments indicate that learning is most efficient when the size of step or length of the presentation prior to review or practice is progressively increased from the earlier to the later stages of the learning task. This proposition, however, is contingent on the amount of positive transfer from each trial to the succeeding trials. As more and more is learned the less room is left for improvement. As the room for improvement gradually decreases the greater can be the increase in size of step. If, on the other hand, the learner fails to notice the crucial cues or fails to make the correct responses to them during the presentation, the greater will be the contribution of periods of active review and practice. One way to make certain that students do, in fact, notice and respond to crucial cues during a presentation is to make provisions for, and allow time for, overt or explicit covert responding.

A half dozen experiments have been reported on the relation between size of steps in linear programs and posttest scores (38, 53, 70, 90, 183, 207). In all except two of these experiments (187, 207) the ones with smaller sized steps produced better posttest scores than the versions with larger steps, regardless of how size of step was measured. In all of the experiments, the response mode was overt. No experiment has been reported in which size of step has been covaried with response mode.

Optimal size of step, however, is related to other conditions on which the contributions of overt responding depend. One is the nature of the learning task. For example, overt responding is an important condition for the acquisition of abstract concepts but not for verbal association learning. In the program used by Evans, Glaser and Homme (38) for converting numbers to a base other than ten, the task was to learn a mathematical principle. This required giving a number and variety of examples. The same is true of the program on learning sets used by Gagne and Bassler (68), and the programs on decimals and squaring two digit numbers ending in five, used by Hamilton and Porteus (90). In these programs each example was a short step. The greater the number and variety of examples, the better was the learning of the terminal behavior. Hamilton and Porteus (89) also used a program on the geography of China which required associative verbal learning to which overt responding usually contributes little. The learning from this program was the poorest of the three that were used. They also found that students with high I.Q.'s learned more and tended to retain more from all three programs than did students with medium and low I.Q.'s. This was probably because the high I.Q. students had higher level of entering ability for each of the programs rather than because they profited more than low I.Q.'s from overt responding.

3. Relation of Response Modes to Time Variables.

The major time variables that may have an effect on the contributions to learning of overt responding are: (a) the reading, observing, or exposure time (b) time to make overt responses (c) time to consult the presentation of the correct answers. The total time of these three is commonly called presentation time or pacing. The time allotted for each, or for the total may be: (a) fixed by the producer of the instructional materials, as it is for most instructional films, ETV presentations, lectures, and tape recordings; or (b) experimentally varied, as it can be when the materials are presented on cards, slides, and by some teaching machines; or (c) determined by each individual learner, as it is when programs are presented in booklets and by some teaching machines.

(a) Rate of development of instructional films. When the amounts of information presented per unit of time by instructional films exceed a certain limit, overt responding becomes almost impossible and may even interfere with learning. Three reference experiments on this point are those of Ash and Carlton (12) on note taking, Roshal (178) on knot tying, and Jaspen (103) on assembling the parts of the breech block of an anti-aircraft gun. Jaspen provided one group of subjects with the parts to be assembled and instructed the members to imitate each step as it appeared on the screen. A control group was not so provided nor so instructed. The experimental group was faced with the problem of shifting attention back and forth from the screen to the performance. Their responses were first observational (covert) and then operational (overt). The responses of the control group were only observational.

Jaspen varied the rate of his film's running time from 7 to 11 minutes, and the level of verbalization from 142 to 45 words per minute.

When the slowest rate of presentation was combined with the highest rate of verbalization, the group that saw this version obtained a mean score on the immediate posttest which was just about double that of the group who saw the same version but were not instructed to imitate the action on the screen nor provided with the materials for doing so. But when the rate of film development was speeded up and the level of verbalization reduced to the minimum of 45 words per minute, the mean test scores of both the experimental and control groups were drastically reduced and that of the experimental (the participating group) was slightly lower. This would seem to indicate that instruction to imitate the action on a fast moving film tended to interfere with learning, and also that rapid movement plus low level of verbalization tended to interfere with effective covert responding.

Roshal (178) provided each of his experimental subjects with a piece of rope and instructed them to imitate the tying movements as they appeared on the screen. He did not vary the rate of development of the film. He found that active participation during showing added very little in learning to tie knots. Relatively few of his subjects succeeded in tying the knots during the presentation.

Three treatments were employed on each of two films by Ash and Carlton (12): (1) students saw the film and took a posttest (2) others were instructed to take notes while viewing the film (3) the last was the same as (2) except that the group was given 10 minutes to study their notes before taking the posttest. For each film a control group took the test without having seen either film. Each film was about 20 minutes long. One was on high altitude flying and the other on ocean survival and safety. The projection room was dimly lighted for notetaking. The criterion test contained about 56 multiple-choice items. Results showed that (1) all students who saw the films under the various experimental conditions averaged about twice as many correct answers on the posttest as the control group who did not see either film; (2) for the high altitude flying film, the students who took notes averaged about one percentage point less than those who did not, and for the ocean survival film their average was about 11 percentage points below those who saw the film but didn't take notes. The obvious conclusion was that note taking actually interfered with learning, probably because of a division of attention. The rate of development of these two films is not reported, yet it is safe to assume that they ran much too fast to permit any gains in learning from the instructions to take notes.

(b) Experimentally controlled rates of presentation. When the material to be learned is divided into small discrete units, as in frames of a program or paired-associates, each item can be presented on a card, a slide or a machine which can be set to move at a fixed rate. Lengths of exposure time can be covaried with response instructions. This was done in an experiment by McGuire (153, ch. 27) who presented a series of nine slides at two rates - 4 sec. per slide (slow), and 2 sec. per slide (fast). Each slide showed a picture of a part of an automobile fuel pump and a name for it. Each name was a five letter word not descriptive of the part. After each sixth time through the list a test trial was given in which only the part was presented and the subjects were required to write its name on a worksheet. Twenty-four male subjects were assigned to each of the two rates of presentation. Each of the two groups was subdivided into three groups of eight men each. One group was instructed to write down the name of each part as it had appeared on the screen during the training series (the overt group), a second group was instructed to "think" or silently rehearse the name of each part (covert group), and the third was simply told that they would see the series six times. All groups were told that there would be a test following six learning trials.

The results in terms of inverse sine transformation of raw scores were:

	Slow Rate	Fast Rate	Difference
Overt	7.8	3.7	4.1
Covert	7.6	6.4	1.2
No. Instructions	6.1	5.0	1.1

Analyses of variance of these data indicate: (a) a highly significant interaction effect between rate and response mode, at the slow rate overt is significantly greater than covert; (b) a similar interaction effect between overt and no participation instructions, the covert being better at the slow rate. In fact the slow rate favored all three response mode conditions. Fast overt responding appears to have interfered with learning much more than covert or no instructions. For the overts the difference between the slow and fast rate is 41 points, whereas for the coverts it is 12, and for the no instruction, 11 points. A comparison between the coverts and the no instruction groups at both rates shows that the gains of the coverts over the no instruction group is almost the same, about 15 points. This would seem to indicate that instruction to think had a motivating effect. Failure to instruct a subject to think or rehearse the correct name does not preclude or prevent covert response. Both groups responded covertly but one more vigorously than the other.

In each of the above experiments the criterion was number of correct responses on a post-learning test. In most laboratory experiments on paired-associate learning the criterion is either number of trials taken to reach a criterion of one or two perfect recitations, or the total time required to reach this criterion level. Using this criterion, Bugelski (24) and Carrol and Burke (30) found that in the learning of paired-associates nonsense syllables by the anticipation method (overt responding) the shorter the exposure time, the greater the number of errors per trial, and also the greater the number of trials to reach the criterion. The longer the exposure time, the fewer the trials required to reach the criterion. Thus the total time required to reach the criterion is about the same in each case. However, as Carrol and Burke noted, Bulgelski's slowest presentation rate was 6-sec. and his materials had low levels of meaningfulness. Materials of high levels of meaningfulness appear to be learned more efficiently at faster rates of presentation.

(c) Fixed vs. Self-pacing. One of the advantages claimed for programmed instruction is that each student can proceed at his own rate. This allows for individual differences in the speed of learning. However, when fixed rates are compared with self-set rates the results fail to confirm this supposition, provided, that the rate is fixed at some optimal level -- not too slow and not too fast.

Briggs (15) compared fixed with self-pacing in teaching airmen to associate names of men with line drawings of twenty real or fictitious electronic tubes by the use of his "Subject-Matter-Trainee." One feature of this machine is that the subject-matter can be either self-paced or fixed-paced. The teaching procedure was in two phases, a presentation phase, and a practice phase. During the first phase, all subjects worked at a fixed pace of 13-sec. per item; on the second phase one group worked at the fixed rate of 13-sec. per item, and each member of another group was instructed to work at his own pace. The first phase consisted of one presentation of the list in which both the stimulus and response terms were exposed. (This was called the prompting trial.) In the second phase all subjects were given 13 minutes for practice. When a stimulus appeared the subject was required to pick the correct response from a display of all 20 of them. An error caused a buzzer to sound and a red light to flash which was followed by a green light indicating the correct response. The criterion test was the number of correct responses chosen from a scrambled order of items and without knowledge of results. The mean test trial scores of the fixed-paced and self-paced groups were almost identical. Also the practice trial scores were not significantly different.

Externally fixed pacing and self-pacing were covaried with response mode and other variables in experiments using linear programs on teaching machines by Feldhusen and Birt (57), and a linear program on basic electricity by Alter and Silverman (4). There were no significant interactions between response mode and pacing mode in either of these experiments.

Carpenter and Greenhill (29) report an experiment in which external pacing was compared with self-pacing in the presentation of 15 programs consisting of 2,055 frames on contemporary college algebra. The same materials were presented in four ways: (1) by a teaching machine self-paced, responses overt (2) programmed textbook, self-paced (3) filmstrip, externally paced (4) conventional classroom teaching. On the final tests there was no significant difference in the mean scores of the externally paced and the self-paced classes. In another experiment using the same subject matter, no significant differences were found between externally paced and self-paced rates even when the former were set at 20% below the mean self-paced rate, or 10% above this mean.

The foregoing experiments cast some doubts on the value of self-pacing in learning from TV, teaching machines or programmed text. Self-pacing, however, has been found superior to externally fixed pacing in the case of reading (60) and in the case of heterogeneous groups (62). Another factor that could influence results of experiments on this problem is individual difference in work and study habits. Some students like to take their time. Working on time schedules could be distracting and interfere with learning. Other students may have developed habits of working faster.

The effect of individual work habits on self-pacing has been investigated by Gropper and Kress (84). In one experiment, in which two science programs were used, it was found that some high I.Q. students worked at rates too slow for efficient learning and others worked too fast for effective learning. But the high I.Q. students, as would be expected, had higher posttest average scores than groups of low I.Q. students. The low I.Q. students who worked at slower rates scored higher on posttest than those working at faster rates. Evidence that these self-adopted rates of work reflect previously acquired work habits is provided by a positive correlation of .80 between individual work rates on two different science programs. The correlation between individual error rates on the two programs was .78. Individual work rates appear to be independent of I.Q. Because these rates are controlled largely by previously established habits they are not, for some students, optimal for either efficient or effective learning.

(d) Relation of response mode to S-R time intervals. When pacing is externally fixed, the time interval between frames in programmed instruction and between stimulus and response terms in paired-associate learning may be experimentally varied as indicated in the preceding section. In paired-associate learning, which is similar in many respects to linear programs, it is possible to vary (a) the time that the stimulus term is exposed (b) the interval before the response term appears, and (c) the time the response term is exposed. A number of experiments have been reported in which the effects on learning of these time variables have been investigated.

In externally paced programmed instruction or in paired-associate learning the order of events is usually: (a) the stimulus term is presented, (b) the student writes his response, and (c) the confirmation or correction (feedback) term is presented.

A fixed amount of time is allowed before the next stimulus term appears. This procedure has been designated as the confirmation method (34). An alternate method is: (a) stimulus term on briefly, then off briefly (b) the correct response term is on briefly (c) the student copies it or echoes it. This is called the prompting method. In a series of experiments by Cook (34), Cook and Kendler (35), Cook and Spitzer (36), it was found that the prompting is superior to the confirmation method in paired-associate learning. In one experiment Cook and Spitzer (36) covaried presentation method with response mode (overt vs. covert) in a 2x2 factorial design. Under the prompting method, one group of subjects was instructed to copy the response term when it appeared (overt responding), and another group was not so instructed (covert responding). Under the confirmation method, these same two conditions were employed. The results indicated quite clearly that the best learning was obtained under the prompting-covert responding condition, and the poorest under the confirmation-overt responding condition.

This result is explained by the authors in terms of the time relation between the appearance of the stimulus term and the response term. According to the principle of temporal contiguity, the shorter this time span, the better should be the learning.

The learning task in this series of experiments was to associate each one of the first ten letters of the alphabet with a geometric code consisting of two lines connecting three dots in a pattern of seven dots arranged in an irregular order. The time required to go through the list of ten pairs once was 72.5 seconds. After each time through the list, a test trial was given. One score was the number of items correct, and another score was the number of incorrect but "legitimate" responses. A legitimate response was one that was on the list of the ten used in the experiment. When the results of the experiment were assessed in terms of different "legitimate" responses, it was found that overt responding had a deleterious effect under the confirmation condition, but not under the prompting condition. This was attributed to the fact that while the stimulus terms (letter) were familiar to the subjects, the response terms (codes) were not. Furthermore, the response terms were selected from a much larger population of possible lines between three of seven dots. Part of the learning task was to discriminate between codes that were and were not on the list. This discrimination was facilitated under prompting by having the subject copy the code when it appeared on the screen. Therefore, on the test trials subjects who copied each of the response terms made fewer "substitute" errors (response not on the list) than subjects who did not copy them. This result is in line with results from other experiments which indicate that overt responding tends to favor response learning more than it does associative learning. As noted above, when the test score was number of correct responses, overt responding was inferior to covert responding under the prompting condition.

The pattern of timing was different in the two procedures. For prompting, the order of events was: stimulus on for 1 sec.; black screen .25 sec.; code line (response term) on for 2 secs.; white screen, 3 secs., during which time subjects wrote the response term from memory; and then black screen for 1 sec. before next stimulus appeared. For confirmation, the order was: stimulus term for 1 sec.; white screen for 3 sec. for subjects to guess the response term; response term on for 2 sec. (confirmation or correction); and the black screen for 1 sec. before next stimulus appeared. Note that the time interval from the Stimulus to the Response terms under prompting was .25 of a sec., while under confirmation it was 3.25 sec. The superiority of prompting was attributed largely to this shorter time interval.

Black (14) noted that in the experiment of Cook and Spitzer, as well as in the previous experiments by Cook (33) and Cook and Kendler (35), that the time interval between the disappearance of the response term on one trial and the appearance of the next trial stimulus was 3 sec. longer for prompting than for confirmation, while the S-R interval was 3 sec. shorter.

Using the same materials as had been used by Cook and Spitzer (36), Black (14) covaried the S-R and the inter-trial intervals both for overt and covert responding, using only the confirmation technique. The stimulus term was on for 1 sec., then subject was instructed to guess the correct response. The overt group recorded their guesses, the covert group did not. After a predetermined interval a signal was given as to whether the R was correct or incorrect. The time delay before giving the next stimulus term was varied. The results showed no significant difference either in respect to overtness or covertness of responding or between longer and short S-R intervals; but there was a significant difference in favor of the longer time delays between trials. Black concludes: "the advantage attributed by Cook and Spitzer to shorter inter-stimulus intervals seems to have been primarily a function of the lengthened inter-trial interval." The inter-trial interval for prompting used by Cook and Spitzer was 4 sec. for prompting and 1 sec. for confirmation.

Another variable that could have favored the prompting method was the fact that an unprompted test trial was inserted after each three training trials. As Lumsdaine (128) pointed out, testing trials are also a learning trial (but without confirmation feedback). At any rate the prompting conditions consisted of three-quarters prompted trials and one-quarter unprompted trials. The learning value of the unprompted trials is that the learner gets practice in testing his knowledge without the aid of a prompt. The question arises as to whether or not there is some optimum combination of prompted and unprompted trials.

Using the same stimulus materials as Cook employed, Angell and Lumsdaine (8) compared "pure prompting," i.e., 12 trials before a test trial, with a mixture of prompting and confirmation trials. This was done for a total of 36 trials with 3 test trials, one after each 12 training trials. They found that on each of the three test trials the groups that got the combination of prompting and confirmation in the ratio of three to one had a consistently higher percentage of correct responses on each of the three tests. The results were interpreted to mean that a sequence of prompted and unprompted trials provides the subject with an opportunity to practice the criterion test — which is never prompted. Lumsdaine interpreted this experiment to support his hypothesis that "vanishing" is a prime condition of learning. The learner must be progressively weaned from relying on prompts and learn to act on his own because the cues have been reduced to a minimum value. But in order to manipulate vanishing experimentally a record of correct and incorrect responses, as the cues are reduced, is essential.

(e) Rate of reduction of prompts. The Angell-Lumsdaine experiment raises the question as to the optimal rate at which prompts can be "vanished." A further question is whether or not the rate of vanishing is related to response modes.

Suppose that this experiment is repeated with all responses covert. What would be the effects on the posttest of vanishing? The answer could depend, to a degree at least, on how well a low error rate is maintained. Experiments cited earlier in this section in which overt and covert response modes are compared have found, even in well designed programs, and when the error rate is low, that the differences in posttest mean scores between the overt and covert modes are not significant. If the prompts are gradually reduced, as they are in well constructed programs so that the error rate remains low from the beginning to the end, one would expect to find that the covert mode is just as effective and the overt one. But if the prompts are faded too fast, there might come a point where overt responding would become the optimal mode. This possibility has not been investigated.

IV. POST-RESPONSE EVENTS

1. Introduction.

In the preceding sections of this paper response modes were considered from the standpoint of events that occur prior to and during the answering or recitation phase of the instructional process. This section deals with events that occur following the answering phase. Such events are variously designated as knowledge of results (KR), or knowledge of correct results (KCR). The general term for all such post-response events is "feedback."

All forms of feedback are stimuli. Some are presented externally and some are produced by the learner's own responses. Example of externally delivered stimuli are: the announcement of "right" or "wrong," an oral or visual presentation of the correct response, the flash of a green or red light on a panel board, and in some instances an explanation of why a wrong answer is wrong. Stimuli produced by the learner include written answers, hearing himself utter a response, proprioceptive feedback from muscular actions, and subjective feelings about the certainty or uncertainty of the correctness of his responses.

Whatever may be the sources of post-response stimuli, their effects on learning depend (a) on the nature of the answering responses, whether overt, covert, constructive or multi-choice, and (b) on how they are interpreted and responded to by each learner. The same KCR may elicit different covert responses depending in part on whether the answering response was correct or incorrect.

2. Relation of KCR to Answering Responses Modes.

Is the effectiveness of KCR in any way dependent upon whether the answering responses were constructive, multiple choice or covert? The experimental literature on this relationship suggests that KCR is more effective when the answering responses are multiple-choice than when they are constructive or covert. Only one experiment has been reported in which KCR versus no-KCR were covaried with overt written responses *vs*. covert thinking responses. It was done by Michael and Maccoby (162) who used an instructional film which was presented in segments with participation sessions interspersed between each segment. The participation questions were administered orally and 10 groups of subjects were instructed to write answers and eight groups were told just to "think" the answers. One half of each of these groups was given immediate KCR after each answer, and one half were not. Control groups were shown the entire film with no interspaced participation sessions. On a 30-item posttest consisting of open ended questions there were no significant differences between the mean scores of the overt and covert groups.

But the groups who received KCR scored significantly higher than the groups who did not. The gains due to KCR were almost exactly the same for the overt and the covert groups. This result held for both high and low I.Q. students.

A half dozen other experiments involving KCR versus no-KCR have been reported. In each of these experiments all answering responses were overt-constructive. The posttest scores showed no significant difference between the mean score of students who received KCR and those who did not. These experiments were reported by: Holland (96), who used sections of the Holland-Skinner program on psychology; Hough and Revsin (101) using a 555-frame program on the history of education; Moore and Smith (165) who used programs varying from 32 tp 39 items on spelling and administered by a teaching machine; McDonald and Allen (142) with a program designed for the teaching of an unfamiliar game; Feldhusen and Birt (57) who used a variety of linear programs presented both on machines and in booklets; Silberman, Melarango, and Coulson (198) with a linear program on simple and compound statements, connectives and arguments; and Krumboltz and Weisman (121) using a 177-frame program on descriptive statistics.

One experiment, however, has been reported by Meyer in which immediate confirmation of correct answers produced slightly better posttest scores than no confirmation (161). The answering responses were written. She used programmed booklets designed to teach 26 prefixes, of Latin or Greek origin, to common English words. The booklets that were distributed to one group contained no KCR, but those distributed to two other groups did. These groups were instructed to put an "X" after each wrong answer, and self-scored their booklet. On the rather difficult posttest the mean scores of the groups who received KCR were slightly better (.06) than those who did not, but on gains from a pretest to the posttest the difference favoring the groups who received KCR was significant at the .03 level. The number of pupils in these three groups were 16, 15 and 14.

The experimental results, such as they are, would seem to indicate than when the answering responses are overt-written, confirmation or KCR contributes little or nothing to posttest scores.

There is some evidence, but again not very conclusive, that when the answering responses are multiple choice, KCR immediately given do seem to improve posttest scores. Angell (10) found that college students who received immediate knowledge of quiz results while doing a program on chemistry that was designed for multiple-choice punchboard responding scored higher on the final examination than students who got the feedback information on the day following the quizzes. Vicory (225) used a 45-minute presentation on number systems with 31 questions to be answered. One group of college students responded by pressing one of a row of buttons and received immediately a signal indicating whether it was the right or wrong one. Another group answered by marking multiple-choice items on an answer sheet and receiving no KR. The group who received immediate KR scored significantly higher on the posttest. In another experiment on number system concepts Vicory and Corrigan (226) repeated the experiment of Cicory and found that multiple-choice responding by button pressing with KR resulted in better learning than multiple-responses on IBM sheets, with no KCR. Kaess and Zeaman (106) used a 30-frame program of psychology vocabulary with multiple-choice responding on a Pressey punchboard. Each of six different groups of students went through the program five times. On the first and subsequent trials two groups had five choices, one with and one without confirmation. On the remaining four trials both groups received confirmation but the one who received it on the first trial outperformed the other groups on all trials except the last one. The purpose of this experiment was to test the effects of "negative information" provided by four wrong alternatives on the first trial. The superiority of confirmation over no confirmation was a by-product.

3. Relation of KCR to the Kinds of Responses Made to Them.

As noted in the introduction to this section, all forms of feedback are stimuli which are either externally presented, response produced or both. The effects that such stimuli have on learning depend on the kinds of responses that they elicit. If they are ignored, as they sometimes are, they would not be expected to have any effect on posttest scores. The elicited responses are first covert, which may or may not be supplemented with overt responses of copying or repeating the correct answers. Like all covert responses their occurrence and nature must be inferred from antecedent conditions and observable changes in behavior. Lumsdaine (130) has noted that the effects of feedback on verbal learning has not been systematically investigated despite the large amount of work that has been done on reinforcement in animal learning. The problem is one of identifying and assessing the contributions to learning of the motivational, affective, and incentive effects on the one hand; and the cognitive, mediational and cue productng effects on the other. It is now possible, however, to identify crudely four or five possible kinds of effects that feedback could have on learning.

If the responses are correct and made to the critical content and not merely to the prompts, confirmations could provide: (a) a rewarding, self-congratulatory, ego-inflating and reinforcing effect; (b) an encouraging, motivating effect to work harder and longer and prevent attention from wandering from the task at hand; (c) an informational effect depending on the extent to which the information supplied can be used either on the posttest or for subsequent learning;

(d) an opportunity to get in some additional covert rehearsal or practice trials if time enough is allowed; and (e) proprioceptive feedback, which for some learning tasks could provide cues that mediate correct posttest responses provided that the stimuli which are fed back are distinctive and unique.

For responses which are incorrect, KCR could have somewhat different effects. Wrong responses sometimes have to be unlearned or extinguished before they can be supplanted by correct ones. This is one reason why programmers strive to keep error rates low. Too many wrong responses, especially in sequence, could be discouraging and motivate the learner to escape from the task. In Meyer's experiment (161), mentioned above, the students were not allowed to score their wrong answers immediately after seeing the correct ones. After the lesson was finished one group was asked to go back over their papers and put an "X" after each wrong answer. These booklets were later checked by the experimenter. Another group did the same and in addition was required to write in a new answer for each wrong one and then score their booklets. She found that the scores which the first group gave themselves contained 8% wrong answers which had not been marked with an "X," but the corresponding percentage of the groups who wrote in new answers was 21%. She observed that to mark an answer wrong is punishing enough, but to write in another and mark it as wrong was doubly punishing. There were no doubt individual differences in these punishing effects.

(a) Reinforcing effects. If a reinforcer is defined as any post-response event which increases the probability that the same or a similar stimulus will elicit the same response, then all of the possible effects of such events on learning, as listed above, would be considered as reinforcers. But if reinforcers are defined as rewards, as they are in operant conditioning, the question arises as to what kinds of events or stimuli are considered by the learner to be rewarding. To what extent is KCR perceived as a reward, by what kinds of learners, and under what conditions? This is a subjective matter in which there are, no doubt, wide individual differences. One way to find out how rewarding KCR is to a learner would be to make it difficult to obtain and thus discover how much effort an individual would exert to get it. Instead of printing correct answers to the frames of a program on the back side of the pages of booklets, or presenting them by merely pulling a lever or punching a button, they could be made available but more difficult to find. An analogous situation is that of looking up words in a dictionary. This is a possible way of measuring the rewarding effects of KCR which has not, to the knowledge of this writer, been investigated in connection with verbal learning.

The reinforcing power of post-response stimuli, whatever it may be, is no doubt acquired by past experiences. For example, the words "right" and "correct" have gained reinforcing power by previous associations with more tangible rewards. A green light on a panel board indicating a correct response is probably better than a red light because green lights are commonly associated with "go" and red with "stop." Acquired reinforcers have different strengths for different individuals because of differences in the degree of past associations with tangible rewards.

Acquired verbal reinforcers may be self-administered. A learner can tell himself that his response is correct. If he is certain that it is correct, then to be told so by an answer sheet or by a green light will probably add little or nothing to the strength of the S-R connection. But if he is uncertain about the correctness of his response, an external confirmation should be reinforcing because it removes the annoyance of uncertainty.

Programs with low error rates were employed in the experiments listed above in which no significant differences in posttest scores were found between groups receiving and not receiving KCR. These programs are so well prompted that most of the subjects answered most of the frames correctly. If a student is correct and knows it, then to find out immediately that he was indeed correct is probably not much of a reward. If, however, a student is highly motivated to achieve a good score and has a fairly high level of test anxiety, KCR may be quite reinforcing. Evidence for this is found in an experiment by Knight and Sassenrath (114) who pretested students for achievement motivation and test anxiety. On a series of linear programs on the construction and an analysis of educational achievement tests which required about two hours per week of work over a period of three weeks, the subgroup which had both high achievement motivation and high test anxiety scores worked significantly faster, made fewer program errors and had a significantly higher mean posttest score than a subgroup which was low on both achievement motivation and test anxiety. The author suggests that one possible explanation for this result is that the highly motivated and anxious students derived greater benefit from immediate KCR. Campeau (27) divided 5th grade boys and girls separately into extremely high and extremely low test anxiety groups and gave a random half of each group KCR and the other half no KCR. The program used consisted of 193 frames on geography. A 54-item posttest was given immediately following the program and again after a delay of 19 days. Of the eight subgroups (four of girls and four of boys) the high test anxiety girls who received KCR scored significantly higher on both the immediate and delayed tests than did any of the other seven subgroups. However, the high test anxiety girls suffered a greater retention loss than the low test anxiety girls. For boys this trend was reversed.

As noted above, multiple-choice answering responses were employed in most of the experiments in which KCR proved superior to no-KCR. In selecting the correct alternative the student is faced with the difficulty of discriminating it from plausible looking incorrect alternatives. In this case students are apt to be less certain of the correctness of their responses than they are of constructed answering responses. To the extent to which this is true, KCR would be expected to have a reinforcing effect.

In experiments on animal learning where cognitive and emotional effects of post-response events are presumably at a minimum, the effects of rewards on learning are a function of (a) the amount of the reward (b) its delay, and (c) the frequency and distribution of rewarded trials, called "schedules of reinforcement." The question arises as to whether or not these variables will exert the same powerful effects on programmed human learning as they do on animal learning.

Krumboldtz and Weisman (118) varied the schedules of reinforcement on a 177-frame program on descriptive statistics and on the interpretation of results of educational tests. The subjects were college students. For one group of subjects 100% of the frames were presented with KCR corresponding to a schedule of continuous reinforcement; a second group got two-thirds of the frames reinforced in a fixed ratio of two out of each three frames being reinforced; and a third group had a fixed ratio of one out of every three; and for a fourth group none of the frames was reinforced. Two other groups got varied reinforcement in the ratios of two out of three frames and one out of three.

The results are reported both in terms of learning errors and number of items correct on a 50-item completion posttest. The greater the number of frames that were reinforced, the fewer were the learning errors. But on the criterion posttest the differences between the mean scores of the four groups who received fixed ratio treatments were small and not significant. For the treatments in which the ratios of reinforcement were varied, no significant differences were found either in learning errors or between mean scores on the posttest. This result could perhaps be attributed to the fact that the frames in this program were well prompted and the error rate was low. The learning prior to post-response events had been so effective as to leave little room for improvement by varying the schedules of reinforcement.

Concerning the effects of delayed KCR, Susan Meyer (161) found that eighth grade students who received immediate KCR and corrected their own papers performed better on the posttest than students to whom the feedback information was delayed until the experimenter had corrected their papers. A similar result was obtained from an experiment in which immediate KCR on quizzes in chemistry was given to one group and delayed for another group until their papers had been corrected (10).

Fleming (59) used two instructional films. One was shown without interruption and a posttest immediately followed. One group received immediate KCR after each answer while another group waited until the next day for their papers to be corrected. A retention test was given later. The other film was shown in sequences with a question and answer period after each sequence. One group received immediate KCR while another waited until the following day for KCR. Immediate feedback produced better learning from both films than delayed feedback. But on a retention test, which was a different form of the posttest, no significant differences were found between immediate and delayed KCR. However, the decrease in retention scores was less for students who saw the sequenced film than for those who saw an entire film. In three schools the feedback was given individually, while in one school it was given orally to the entire class. For some reason the students in school where it was given to the entire class, scored higher both on the posttest and the retention test.

The evidence, such as it is, indicates that attempts to increase the amount of reward for correct responding have not been successful. Moore and Smith (164) using the Holland-Skinner program on psychology with college students gave one group of subjects a penny for each correct response in addition to KCR. The mean posttest score for this group was no higher than that obtained by a comparable group who received only KCR and another group who received no KCR. There is some evidence that KCR is more effective than KR, i.e., the announcement of right or wrong. Briggs (16) employed several methods of training by the use of a device called a "subject-matter trainer" and found that KCR given immediately after each response produced better learning than a mere right or wrong indication. Hirsch (93) also found that KCR was more effective than KR and that repeating the question and the answer was even better.

Vicarious Reinforcement

In classroom discussions, students are in a position to observe whatever positive or negative reinforcements may be administered by the teacher to their classmates. The question arises as to whether the results of such observations have a reinforcing effect, in the sense of strengthening the associations between specific stimuli and responses, or have only an informational effect. The only thing a student may learn from such observations is the kind of behavior that is rewarded or punished.

Sechrest (186) has reported an experiment in which young children worked in pairs, each on a different jigsaw puzzle. When one child finished his first puzzle, the experimenter would either give positive praise, or a faint one and in the presence of the other child. Then a second puzzle was presented. The problem was to determine the effect on (a) the child who got the reinforcement, and (b) the child who did not. The results show that it had a positive effect on the child who received reinforcement but had no effect on the observer child.

Van Wagenen and Travers (224) obtained a similar result. The task was to learn German vocabulary. Groups of eight children at a time were presented with a German word and two English words. The teacher would call on different ones to give the correct word. Four of the eight were selected as respondents and each got five out of 20 words for four trials. The results obtained from this situation were compared with results obtained when students worked on the same materials in isolation with teaching machines. Under one condition the feedback was supplied by the machines and under another by the teacher. On the criterion test there were no significant differences either between direct and vicarious reinforcement, or between the teaching machine and the classroom situations.

(b) Motivational effects. In addition to whatever reinforcing values KCR may have, it also may have certain incentive or motivating effects. In working through a well-prompted program a student who finds that his answers are correct time after time may feel that he is getting somewhere, making progress toward a goal. This may inspire him to work longer and harder. One of the advantages claimed for teaching machines over programmed textbooks is that operating a machine has certain game-like qualities. The student "plays" the machine. Skinner (201) reports that many students said that machine work was fun and challenging.

People generally prefer to do the things they enjoy doing and avoid or postpone doing things they dislike. But the things that people like to do are apt to be the very ones they have learned to do well. Learning comes from trying new and challenging tasks. When a printed program or a teaching machine loses its challenge the work is likely to become boresome and tiring. This is particularly true of long programs of a thousand or more frames. But students who report that programs are boring may be the very ones who also find textbooks and other instructional activities boring.

Several studies have been reported on student attitude toward programmed instruction. The results indicate that attitudes are not related to amounts learned. Eigen (48) compared attitudes of students toward a Skinner-type teaching machine and programmed textbooks and found that the text were liked the better. There was, however, no significant posttest difference between groups having positive or negative attitudes. Likewise, Feldhusen and Eigen (58) found low correlations between amounts learned from a program and students' attitude toward it. But for students in the eleventh grade the correlations were more positive and higher than for students in the ninth grade. Studies of the relation between liking and learning from instructional films and ETV programs indicate low correlations between these two variables (150). There is need for a more detailed analysis of the aspects of programs that are liked or disliked. Do students who like a program do so because of immediate KCR, or because of self-pacing or low error rates? Do they dislike it because it is monotonous, or because it is too hard and they make too many mistakes?

The announcement of "wrong" or the flash of a red light to indicate that a mistake has been made is called a negative reinforcer. The effects of negative reinforcers on learning has been investigated by Melarango (160). Ten instructional frames were prepared for each of five symbols employed in symbolic logic. Five ambiguous items, which looked legitimate were inserted at various points among the instructional items. Each frame was presented on a card and exposed through a window. The subjects responded by pressing one of a row of buttons. The experimenter flashed a green light if the response was correct and a red light if it was incorrect. The joker items were signaled as correct for one group of subjects and incorrect for two other groups. The experiment was designed so that one of the groups of subjects got 55 positive and no negative reinforcement; another group got 50 positives and 5 negatives with the negative spaced throughout the series; and a third group got 50 positives and 5 massed negatives. In order to achieve this balance the experimenter signaled some correct responses as incorrect when they were in fact correct, and vice versa. The posttest consisted of 45 items. The results indicate that the spaced negative reinforcers did not depress the mean posttest score below that of the group who received all 55 positives; but the massed negative reinforcements interferred significantly with the learning of the symbols which were adjacent to the place in the series where the massed negatives were inserted. This result could be attributed to the demoralizing effect of getting five red lights in succession. One wonders what the result would have been if the answering responses had been written instead of multiple-choice.

The more effortful mass responses are, the more likely they are to build up reactive inhibition which, according to Hull's learning theory, interferes with learning. It accumulates gradually and is dispersed by rest periods. Faison, Rose and Podell (55) found that inserting brief rest periods at intervals during the showing of a 20-minute instructional film had the effect of adding a small, but statistically significant, increment to posttest scores. The difference between the "pause" and "no pause" groups was greater during an afternoon session when the accumulated fatigue of the day may have had an interfering effect. Evidence that the pause did not serve to dissipate reactive inhibition was that the part of the film seen just after the rest periods were no better learned than the parts that preceded the pauses.

A similar result was obtained by Schoer (184) who divided college students into two groups on the basis of a specially devised test for measuring how fast reactive inhibition is generated. Both fast and slow groups were then given a 1,200 frame linear program on statistics and measurement in education. The responses to the frames presumably were written with immediate KCR. Each student did 80 frames a day for 15 days. The posttest consisted of 36 fact items, and 36 application-multiple-choice items. Contrary to expectation based on previous experiments, it was the high or fast inhibition groups who made the fewest errors on both parts of the posttests. No significant posttest differences were found for subgroups who scored high and low on a relevant vocabulary test. This result could have been due, in some part, to the rate at which the member of each group worked through the program. It may be that in programmed instruction that KCR is powerful enough to overcome the negative effects of reaction inhibition.

(c) Informational effects. Knowledge of results can in some cases provide information that a learner can use in answering posttest questions and on subsequent learning trials. For example, in shooting at a target a rifleman gets information by observing how far from the bullseye a shot may be, and use the information to correct his next aim. In this case, the overt act of shooting produced a stimulus that could be compared with correct answers. The value of such comparisons depends on the uses that can be made of the information. KCR provides more information than KR but has no greater effect on posttest scores unless more use can be made of knowing the correction for an incorrect response than merely knowing that it was wrong. In multiple-choice responding if only two alternatives are presented, a wrong choice provides as much information as a right one. This is illustrated by the games of 20 questions where a "yes" or "no" answer will rule out half the population of possible answers. When more than two multiple-choice alternatives are given, Bryan and Rigney (22) found that immediate knowledge of results plus an explanation of why an answer was right or wrong produced significantly better posttest scores than mere immediate KR. In a later experiment by Bryan, Rigney and Van Horne (23), three different types of explanations of right and wrong answers were compared but no one proved to be superior to the others. Krumboltz and Bonawitz (120) found that confirmation presented as a complete sentence was more effective than giving it in a single word or phrase. They used a 153-frame linear program on educational psychology. A similar result was obtained by Hirsch (93) from an experiment in which participation sessions were interspersed in an instructional film. He found that KCR was better than the mere announcement of "right" or "wrong," and the best results were obtained when the correct answers were given as completed sentences. These experiments indicate, but do not prove, that the informational value of feedback depends on the uses that can be made of it either on posttests or in connection with subsequent learning.

Cummings and Goldstein (43) advanced the hypothesis that the more complete and accurate informational feedback is, the greater will be the benefits that the learner can derive from it, up to some limit beyond which it becomes redundant. To test this hypothesis they presented groups of college students with a program on an unfamiliar topic — "The Diagnosis of Myocardial Infarction." One group wrote answers to 119 frames in booklets, while another group was instructed to "think" the answers but not write them down. Some of the required responses were fill-in words and others called for recognition of differences between graphs of electro-cardiograms. The posttest test was composed of similar items and consisted of 150 questions. It was administered immediately after completion of the program and again after a delay of 10 days. Time for completion was recorded. Scores were calculated separately for the pictorial items and the verbal ones.

The test performance for the overt group was significantly better than for the covert on the pictorial section, both on the immediate and delayed posttests. The mean scores of the overt group were also higher on the verbal items, but not as much so as on the pictorial ones. The overt group required an average of 45.5 minutes longer to complete the program than the covert group.

The authors predicted that the overt group would be superior to the covert group on the pictorial items but not on the verbal ones. Perhaps one reason why this group was also superior on verbal items was that in covert responding, which requires the recall of a large amount of new materials, the feedback of information does not provide the learner with a complete record of his responses. It may be that when required responses are longer and more complicated, the value of overt responding is greater because it provides the learner with a record to which his responses can be compared.

Travers, et al (221) report an experiment indicating (a) that a certain degree of redundancy in informational feedback facilitates learning, and (b) that learning is best when the last item in the feedback information is the correct response. The task was learning German vocabulary. A German word was presented and followed by two English words. The task was to name the correct English word. Four types of feedback were employed: (1) if the student gave the correct answer the teacher said "that's right," if the wrong answer was given, the teacher said "that's wrong"; (2) for a right answer, the teacher said nothing, for a wrong one, the teacher said, "wrong"; (3) for a right answer, the teacher said "right," for a wrong answer the teacher told the right answer; (4) if the answer was right, the teacher said nothing, if wrong, the teacher gave the right answer. Conditions (3) and (4) produced the best learning. Condition (2) was the poorest. Thus it would appear that when a correct response is given, being told that it is correct is less effective than being told the correct answer when a wrong one is given. Also knowledge of the correction after a wrong response is more useful than after a right response. One might suppose that where there are only two alternate responses, the knowledge that one is wrong would convey the information that the other must be right. For some students this may have been true, for others it apparently was not. The subjects were 4th, 5th and 6th grade students.

(d) Additional practice trials. In the previously cited experiment by Michael and Maccoby where KCR was found to produce better learning than no-KCR under conditions of either overt or covert responding, the authors noted that the announcement of correct answers not only had a possible reinforcing effect but also gave the students an opportunity to covertly rehearse the corrections for incorrect responses. Also the results obtained from the experiments by Bryan and Rigney (22), Krumboltz and Bonawitz (120), and Hirsch (93) indicating that correct answers given in content with questions are more effective than single words or phrases, could be accounted for, to some extent, by the fact that reading such sentences is indeed an additional covert practice trial.

The principle of contiguity in associative learning requires that the correct response be emitted either in the presence of or immediately following the stimulus term. It will be recalled that in the experiments on prompting versus confirmation Cook and Spitzer (36) attributed this superiority to the fact that the time span between the stimulus and response terms is much shorter than it is under the confirmation procedures. In programmed booklets where the correct answers are given on the back sides of the pages, the questions and the correct answers are not in view at the same time. In other programs the correct answers are printed in the right hand margin of each page and exposed by moving a masking slide.

3. Response-produced cues.

A theoretical advantage of overt responding is that it produces stimuli which may become conditions to correct responses and hence serve to mediate them on retention tests. For example, a written response produces both proprioceptive and visual stimuli. The articulation of a response produces both auditory and proprioceptive cues. When overt responses are correct, knowledge of correctness may function to associate, the response-produced cue, with them. This effect is most clearly demonstrated in perceptual-motor learning of a sequential performance when each act produces stimuli that cue off the next set.

In verbal learning, response-produced stimuli may function as KR. A response may be accepted as correct or incorrect on the grounds that it looks right or wrong, or otherwise "seems" to the learner to be right or wrong. In a highly prompted frame the learner should be able to judge for himself whether his response was right or wrong. One of the reasons why overt responding may sometimes interfere with learning is that response-produced cues are not distinctive enough to become specifically associated with correct responses. The response-produced cues may have previously been associated with correct responses. The response-produced cues may have previously been associated with other responses, or the same cues may become associated with different responses.

Knowledge of correct results may have any one or all of these potential effects on learning. It may be motivating and reinforcing, provide useful knowledge, encourage additional covert answers on posttests. The relative contributions of each of these effects to the learning of different tasks by students of different entering abilities has not been investigated.

V. INDIVIDUAL DIFFERENCES

1. Introduction

When instruction is conceived as a process by which students are moved from imperfect to more perfect performances, individual differences appear in three major dimensions. One is the distance that separates each individual from where he is at the moment to the point of arrival required by the educational objective. If the point of arrival is measured by a perfect score on a reliable and valid posttest, the distance to go is measured by the score on a pretest. A second dimension is differences in the amounts of previously acquired knowledge and skills, not measured by pretests, that can be transferred and used for the attainment of the goal. This repertory of equipment is usually measured by educational achievement tests. The third dimension is differences in abilities to select from this repertory the particular items of information and the relevant cognitive skills and apply them to the learning task. This ability is usually measured by tests of general intelligence. In specific instances it is measured by the rate of learning.

Travers (220) has divided the second dimension mentioned above into two parts (a) prerequisite learnings such as learning to add before learning to multiply, and (b) learning set variables which may facilitate or interfere with new learning. The interference effect is important in cases where previously required responses have to be extinguished or unlearned before new learning can occur.

The question now arises as to the extent to which teaching strategies should be varied in accordance with individual differences in entering abilities. The particular concern of this paper is the extent to which optimal response modes are predetermined by individual differences in entering behavior. The experimental literature, such as it is, suggests few general propositions which merit further investigation.

(a) The more students know about a subject or topic (i.e. the more of the to-be-learned responses that have already been acquired), the less will be the contribution of overt responding to further learning.

This proposition could be tested by discovering (1) whether students who score low on a valid and reliable posttest will profit more from overt responding with KCR then the students who score high, and (2) whether students who have had previous instruction in a subject will gain more knowledge by responding overtly to the frames of a program based on the subject than a comparable group who were not previously instructed in it.

(b) The greater the student's repertory of prerequisite learnings and cognitive skills relevant to the task, the less will be the contributions to further learning by overt responding.

This proposition could be tested by finding out whether students who score high on relevant educational achievement tests will profit less by overt responding to a program than students who score low.

(c) Students with low I.Q.s (below 75) will profit more from overt responding to programs or any other presentation, than students with high I.Q.s (above 125).

Students with high I.Q.s have not only greater repertories of previously acquired relevant knowledge but also greater ability to apply it to new learning tasks.

If these propositions are valid, one would expect to find that: (a) very young children whose repertories of previously acquired knowledge and skills are limited will profit more by overt responding with KCR than older children; (b) the higher a child's mental age or I.Q., the less important is overt responding, and (c) the greater the difficulty of the learning task, the greater will be the contribution to learning of overt responding. Some of the experimental evidence bearing on these predictions will now be presented.

2. Chronological Age

The few experiments on young children in which overt and covert modes have been compared indicate that overt responding with KCR is the superior mode. In a program designed to teach reading to kindergarten children, McNeil (155) found that pronouncing words was definitely superior to merely looking at them. The overt mode was particularly beneficial to children with high I.Q.s. For some reason boys learned more from the program than girls. Suppes and Ginsberg (213) taught concept formation to groups of kindergarten and first-grade children by using programs on the concepts of 4-ness, 5-ness, sets and numbers. Five experiments are reported involving stimulus and response variations. In the first experiment one group of children was required to make an overt correct response after each incorrect one. The members of a control group were merely told whether their responses were correct or incorrect. The experimental groups performed significantly better on the posttest. In this experiment the comparison between overt and covert responding was in relation to post-response stimuli. The answering responses of both groups were overt.

Keislar and McNeil (108) found no significant difference between overt and covert responding on the part of first and third grade children on a 432-frame program which had a 14% error rate. The program was on physical science and the children worked 20 minutes a day for twelve days. The frames were on a sound filmstrip. No reading was required. Children in the overt group indicated their responses by pressing a selector button. If the response was correct a green light flashed; if wrong the experimenter advanced the presentation to the next frame. Children in the covert group were instructed to look at the frame for five seconds after which the correct response was shown by a green light. The posttest was a multiple-choice picture test followed by a standardized interview. The test items were all different from the presentation items.

The results of this experiment are not in agreement with those obtained from the two that are cited above. There are, however, important differences between them. The learning tasks were different, as were the posttests. The most important difference was perhaps the way in which the post-response events were treated. In the Suppes-Ginsberg experiment the "overts" were required to correct their incorrect responses by repeating the correct words in the presence of the stimuli to be learned. The "coverts" were merely told whether their responses were right or wrong. In the Keislar-McNeil experiment this situation was reversed. When the overts made an incorrect response there was no feedback of correct answers; but when the coverts "thought" a response, a green light indicating the correct one was flashed after a period of five seconds. This provided the covert group with an opportunity to rehearse (silently) the correct S-R association, but the "overts" were denied this opportunity. But in the Suppes-Ginsberg experiment it was the "overts" who had not only the opportunity but were required to rehearse the correct S-R associations. Thus it would appear that an effective post-response mode is one that facilitates the rehearsal of to-be-learned associations. This is true regardless of the age of the students.

Silberman and Carter (195) report the results of some exploratory "tutorial" studies designed to improve the effectiveness of programs in reading, arithmetic, geometry and Spanish for students of different grade levels. From the data derived from trying out the programs on one child at a time and noting the difficulties that were encountered, a number of hypotheses were generated about how each program could be improved for children of different ages. In regard to pre-presentational factors, it was found that "first graders became inattentive when even a small amount of verbal direction preceded the task. Prompting them to perform the task was a better procedure than telling them about the task." (p. 81). As to presentational factors, short steps covering every detail of the content and every specific skill to be acquired, with frequent overt responses, was found to be very important for first graders. The most important post-response event for first graders were immediate and frequent tangible reinforcements, and the least important were previews and summaries. On posttests first graders were much less able to answer transfer items correctly than were older children. The greater the departure of items on the posttest from the materials explicitly covered by the program, the lower the posttest scores obtained by younger children. But with older students, such discrepancies can be greater without reductions in scores. These differences in programming techniques for children of different ages seemed to hold for all four programs.

3. Mental Age and I.Q.

The higher the mental age of a child the more likely he is to have acquired the abilities that are needed for the acquisition of new knowledge and abilities. In view of the fact that the benefits derived from overt responding are dependent upon entering abilities, one would expect that children whose mental ages are low would benefit more from overt responding than those of higher mental ages.

For children whose mental ages were above the median of a group of first and second grade children, Wittrock (232) found no significant difference between overt and covert responding to a program designed to teach the relations of molecular action to the phenomena of evaporation and condensation. But for children whose mental ages were below the median, the overt mode was significantly more beneficial. But on a retention test given a year later, no significant difference was found between mean scores of groups who responded overtly or covertly. The program consisted of colored slides made from drawings and accompanied by a tape recording. The task was to learn the word or words associated with each drawing. The "coverts" were not required to respond orally to the blank spaces in each drawing. The posttest consisted of a ten minute standardized interview plus a multiple-choice picture test.

In most experiments in which response mode has been covaried with general intelligence the subjects were divided on the basis of scores on a standardized intelligence or aptitude test. In the classic experiment by Hovland, Lumsdaine and Sheffield (100) on teaching the phonetic alphabet to soldiers, the subjects were divided into high and low ability groups on the basis of the army intelligence test. It was found that active participation in review sessions was more beneficial to subjects of low ability, particularly on the more difficult items, than to those of higher abilities. Overt responding did tend to narrow the gap in posttest mean scores between subjects of higher and lower mental abilities.

This experiment, however, leaves unanswered the question as to whether the learning benefits derived from overt responding by low ability groups was due to increased motivation or to more effective practice. The experiment was repeated later by Lumsdaine and Gladstone (134) with an alternate but less embellished version. The results failed to confirm the previous finding that overt responding is more beneficial to low I.Q. subjects than it is to those of high intellectual abilities.

In both of these experiments the rate of presentation was fixed. An advantage claimed for self-pacing is that it favors the slow learners. A review of the literature on self versus fixed-pacing, as presented earlier in this paper, indicates that the advantages of self-pacing may not be as great as they have been claimed to be. To allow some learners to proceed at their own pace may be a disservice. Brooks (19) found that on frames in a program on which students made errors, the time spent in reading them, or pondering over them, was significantly greater than time spent on frames to which the answers were correct.

A problem of great concern to programmers is how to speed up the learning of slow learners so as to reduce the effects of individual differences in mental abilities. To this end certain program variables have been experimentally manipulated. One is the degree of prompting which controls error rate. Bean (13) used the Henmon-Nelson Test of Mental Ability to divide his experimental groups into high and low ability subgroups. The program was on plane geometry and consisted of 951 linear and 852 branching frames. The response mode of all subjects was overt. On the linear program the error rate of both the high and low ability groups was low and about the same for both groups. On the posttest the mean score of the high ability group was significantly higher than that of the low ability group. This difference could be due to the kind of items used on the posttest. If the geometry problems on the posttest were selected from those used in the frames, the differences in mean scores would have perhaps been less than if the posttest contained only transfer items. But the memory for propositions and proofs is, of course, not a valid test of knowledge of geometry.

Lambert, Miller, and Wiley (122) also divided their subjects into high, medium, and low ability groups on the basis of the Henmon-Nelson test. The program was 843 frames on sets, relations, and functions prepared by Eigen. It was presented in sets of booklets. A posttest was administered at the completion of each booklet and again at the end of the program. Each ability group was subdivided into an overt and covert responding group. The error rate of the program was low as determined by the records of the overt groups. An analysis of the posttest scores indicated a slight, but not significant, interaction between level of intelligence and response mode. For the high ability groups there was no significant difference on the posttest between the overt and the covert responders. But, surprisingly enough, the covert mode was slightly superior to the overt for the medium and low ability groups.

This experiment, however, has been criticized by Holland (1965) on two main counts. By using his black-out technique he found that some of the frames could be answered correctly without reading the critical content. He also noted that the programs were printed on thin paper with correct answers on the back side of each page. He found that of the 329 items in one of the booklets, the answers to 201 could be read without lifting the page, and another 71 could be read by lifting it slightly. This fact could have accounted for the low error rate of the program.

It seems likely that neither self-pacing, nor high level prompting, nor overt responding will suffice to speed up the learning of slow learners to anything near the level of the fast learners. What about other variables?

Angell and Lumsdaine (5) found that partial cueing of correct responses to hard items tended to favor slow learners. The task was to learn the airlines' code for each of several cities. Partial cueing consisted of permitting the student to see one, two or all three of the letters in each code.

There is some evidence that KCR is slightly more beneficial to slow learners than to fast ones. Little (125) found that KCR was more helpful to the slow learners than to faster ones in learning educational psychology. Michael and Maccoby (162) interspersed practice or review sessions in an instructional film. One group engaged in active participation while a control group participated covertly. Both groups received immediate knowledge of results. Each group was subdivided into high and low ability students. The low I.Q. groups who responded covertly profited more from receiving KCR than did the high I.Q. groups who responded covertly. The difference in gains was 3.3 percentage points. But for the groups who responded overtly the difference favoring the low I.Q.s was only .8%. The subjects in this experiment were high school juniors and seniors. The range in their I.Q.s is not given. The mean difference was probably not very large.

There is some evidence that intrinsic or branching programs are better than linear ones for bringing slow learners up to a level of achievement that will approximate that of the fast learners. Branching always requires overt responding as does computer based programming. Coulson and others (39) found that the remedial materials in a branching program did tend to raise the performances of slower learners to that of faster learners, but it took considerably more training time to accomplish this advance. This result, however, is not supported by a number of other experiments in which no important differences on posttest were found when the same materials were programmed linearly and intrinsicly. It may well be that the format of a presentation is not the answer to the problem of individual differences.

The answer may be found in the more careful diagnosis of and adjustments to special abilities. The important differences between slow and fast learners may not be general intellectual abilities, as measured by intelligence tests, but rather specialized entering abilities (52, 71, 105, 211).

Jensen (105) compared retarded, average and gifted junior high school students in respect to rates of learning a simple task of associating the positions of a circular array of buttons on a keyboard with different colored triangles, squares, and circles. All three groups were of the same chronological age (about 14) and their mean I. Q. s were 66. 17, 103. 04, and 142. 54 respectively. For each task five stimuli and five buttons were used. After 200 trials of practice a new task was introduced, up to a limit of six tasks. When a correct button was pressed a green light flashed and remained on for one second before the next stimulus appeared.

When a wrong button was pressed, all other buttons went "dead." The gifted group learned the first task to perfection in about 100 trials. But after 200 trials, the score of the retarded groups was only slightly above that which could be obtained by guessing. On the second task, the retarded group was given extra reinforcement. When the green light flashed, the experimenter would say "good" or "that's right." This raised the asymtote of their learning curve somewhat but not very much. Thereafter the retarded group was given practice in naming each stimulus as it appeared without pressing a button. After each subject could quickly say "blue triangle" or "red circle" etc., the learning task began. On the remaining tasks some of the retarded subjects could learn very well without naming each stimulus before selecting a button. Others could not, but did learn very well when they named each stimulus before selecting a button. On the sixth and final task the retarded group advanced somewhat more slowly than the average and gifted groups, but after 200 trials their performances almost reached the level of the other two groups.

An interesting result of this study was the wide range of individual differences in rate of learning among the members of the retarded group. One of the two fastest learners in the entire study had an I.Q. of 65! This student probably possessed some specialized ability related to this particular task that is not measured by an intelligence test. On some other task this fast learner might be slow. The slow retarded learners were, no doubt, deficient in using verbal mediators in learning, which would account for their improvement when given practice in stimulus labeling.

4. Entering Abilities and Task Difficulty.

With the advent and future promise of individualized instruction by the use of computers, the need for diagnosing and measuring each individual's entering repertory of knowledge and skills relevant to each learning task is of paramount importance (212). At the moment, it is a relatively neglected area of research on programmed instruction (77). One approach to the problem is to consider it from the standpoint of task difficulty. Why is it that some students find a program easy and others find it very hard — even if it is well prompted? Why is it that some students may correctly answer most of the frames of a program and still score low in the posttest? One answer to this question is that some students may learn only to associate the correct answers with the prompts rather than with the critical content. When the prompts are absent on the posttest they are hopelessly lost. This sad state of affairs is most likely to happen with students whose repertories of previously acquired knowledge and skills are deficient in what it takes to learn the critical content effectively.

In order to learn the crucial content of a lesson efficiently, a student must have already acquired the ability: (a) to read or to listen carefully, to notice the key word, to discriminate the relevant from the irrelevant cues; (b) to perceive correctly the meanings of the key words; and (c) to have available in his repertory of responses the ones that are to become associated with the meanings contained in the key words.

These three requirements for efficient learning do not, by any means, exhaust all of the possible aspects of entering behaviors. They will suffice, however, to illustrate the sources of individual differences in task difficulty. Stated in terms of previous learnings they represent stimulus-learning, perceptual-learning and response-learning. Students who are deficient in any one of these abilities and have to learn them while learning the lesson will profit more by overt responding to program frames than will students who are not deficient in any one of them.

(a) Task difficulty due to deficiencies in stimulus discrimination. Most A-V presentations, textbooks and lectures contain (1) critical content to be learned — i.e., relevant cues (2) contextual materials, and (3) illustrations, examples, and motivational embellishments. Programmed materials are usually composed of cues, context, and enrichment materials (113). The inability of a student to discriminate between the relevant and irrelevant content, or between the important and the trivial words is a major handicap to learning (98).

The difficulty may be due to either a lack of familiarity with the materials or inability to discriminate between the relevant and the irrelevant. Inability to discriminate may be overcome by discrimination training which involves the reinforcement of the correct response to a stimulus and the extinction of all incorrect ones by no reinforcement. Discrimination training in laboratory experiments always requires overt responding. If the materials to be learned, either from a program or any other form of presentation, require a fair amount of discrimination learning, the optimal response mode would be expected to be overt. If the difficulty is due to a lack of familiarity with the content of the material, then some pre-presentation familiarization teaching may be required.

The importance of stimulus familiarization is illustrated in an experiment by Wulff and Kraeling (236) on learning to assemble the parts of an automobile ignition distributor. Two procedures were compared. One involved overt pointing to critical features of each of the parts which were mounted on a board and shown one at a time in a predetermined order. First, the experimenter pointed to the feature (or features) of each part that had to be noticed and discriminated from other features in order to assemble it in its proper place. Then each part was handed, one at a time, to each of the subjects who was required to point to its essential features. Immediately after this period of stimulus discrimination training, the film was shown and a performance posttest was given. The test was scored both for selection and assembly errors.

An alternate procedure consisted of presenting close-up shots of each part just before it was shown assembled in the film. While each close-up picture was on the screen the narrator called attention to its important features. In this procedure stimulus discrimination was learned by covert responding, whereas in the pre-film training session it was learned by both covert and overt responding. On the performance posttest the mean number of selection and assembly errors was significantly less for the "covert-overt" groups than for the "covert" only group.

An experiment by Wulff and Emeson (235) indicates still further the importance of stimulus discrimination learning as a prerequisite to associative learning. The task was to learn the names of eight drawings of electric circuits. There were four pairs of two circuits each. The members of each pair were closely alike and required close observation to notice the small difference. The differences between pairs were more obvious. The names were familiar words such as "shaper" and "limiter" for one pair, and "mixer" and "follower" for the two circuits in another pair. On the criterion test the diagrams were presented one at a time on a device called a Subject-Matter-Trainer which exposed all eight of the names. The task was to push a botton under the chosen name. If correct, a green light flashed; if incorrect, a buzzer sounded and the experimenter announced the correct name before advancing to the next diagram. A perfect performance was defined as two correct responses for each pair or a total of 16 correct answers. The subjects were nine Air Force Technicians. After two hours of practice on the machine no one got a perfect score. Two subjects got up to 15 and one got no further along than eight. This demonstrated the need for stimulus discrimination training prior to associative learning.

This training consisted of sorting the eight circuits, printed on cards, into each of eight boxes, numbered 1 to 8. Each box bore the label of a circuit. The task was simply to match the circuits on the cards to the labels on the boxes. This was learned very quickly by all subjects. Then, as an intermediate step toward name-associate learning, the labels on the boxes were covered but the position numbers were still exposed. The task was to put each card in its correctly numbered box. This task also was learned fairly rapidly.

The next step was to teach the association between each circuit and its name. This was done on the Subject-Matter-Trainer. The procedure was the same as before. Each circuit was exposed one at a time and the subjects were required to push the button under the correct name. On the first trial the box number of each circuit appeared over the correct name as a kind of a prompt. If the right name was chosen the green light appeared, if not, a buzzer sounded and the correct name was announced. All subjects attained perfect scores in less than 30 minutes.

This compares with no one getting a perfect score after two hours of work before the stimulus-discrimination training was introduced.

In regard to the response mode used during stimulus training, card sorting was, of course, a form of overt responding. Another group was given discrimination training by being instructed to study the circuit-word cards for 20 minutes rather than sort them into boxes. Each subject was tested after 20 minutes and again after 40 minutes on the trainer. Only two of the seven got perfect scores after 40 minutes of training. Thus it would appear that the overt mode for stimulus discrimination training was the superior one. However, the time spent in card sorting was not reported.

(b) Task difficulty due to deficiencies in response learning.

In commenting on the conditions under which overt responding may be expected to facilitate learning, Lumsdaine and May (135) guessed that one such condition would be when response learning is required in addition to learning associations between a new stimulus and a familiar response. Underwood (222) also voiced the opinion that in paired-associate learning the meaningfulness of the stimulus is apt to be relatively unimportant. This means that it is much more difficult to associate an unfamiliar response to a familiar stimulus, than vice versa. For example, it is easier to learn CVX-DOG than to learn DOG-CVX. If it is true that the greater the difficulty of the learning task, the more effective will be some form of overt responding, one would expect to find that overt responding will facilitate learning the uses of technical terms and learning materials that are strange and unfamiliar. This expectation is supported by some experimental results.

Cummings and Goldstein (43) found that overt responding to programmed materials on the medical subject of Myrocardial Infarctions yielded better learning than covert responding. Williams (230) compared two modes of overt responding, construction and multiple choice, with two modes of covert responding, reading with key word underlined and reading with no underlining. The materials were taken from the Holland-Skinner programs on the Analysis of Behavior. The subjects were college students who used the programs for review after having had previous instruction. The criterion test was 20 items from frames plus two essay questions.

The two groups who were instructed to respond overtly (construction or multiple-choice) had mean scores significantly higher than the groups who were instructed to read. This difference held for students with high and low verbal aptitudes. The objective test contained 20 items of which eight required responding with a technical term. The mean test score on these eight items obtained by the overt-construction group was significantly higher than the mean obtained by either of the other three groups. This result confirmed the hypothesis that when the response term per se, overt practice in writing it out should increase familiarity with it and hence aid learning. This hypothesis might be generalized into the proposition that the greater the number of response words that are infrequently used, as in the Thorndike-Lorge word count, the greater will be the effectiveness of requiring constructive overt responses to frames in a program. The results of an experiment by Eigen and Marguiles (50), cited later, lend some support to this generalization. Additional support is given by the fact that Mrs. Williams' subjects had previously read the book on which the programs are based and therefore had received a degree of familiarity with the technical terms. But despite this fact, the constructive response mode on these terms was still superior to other modes. Had the subjects worked through the programs with no previous knowledge of the content the differences would have been greater.

An indication of the effect of familiarity with content due to previous study is given by the scores of a comparable group of students who took the criterion test without going through the program. The mean posttest score was 8.8 points out of a maximum of 29, while those of the construction, multiple-choice, reading underlined, and reading without underlines, were 23.5, 23.0, 20.7, and 20.6 respectively. A similar comparison on the eight technical items is not reported.

An experiment, cited above, by Meyer (161) on teaching the meaning of words with prefixes derived from Greek and Latin illustrates the difficulty that students have in learning when the to-be-learned response terms are not immediately available. One lesson, for example, contained this frame:

> "A flame emits or sends out light.
>
> A person who has left his native country is an __migrant."

Most students wrote im in the blank space and some refused to believe that there is any such word as "emigrant." On the other hand, students had little difficulty in supplying circum to navigate after it had been explained that circum means "around." The word "circumnavigate" was already in the vocabularies of most of these students and the word "emigrant" was not.

(c) Task difficulty as determined by the degree of meaningfulness of the materials. One of the best established principles of verbal learning is that the more meaningful the material, the less are the benefits to be derived from overt responding. But when the materials are relatively meaningless, overt responding is the optimal mode.

A classical reference experiment is that of Gates (73) who varied the proportions of total learning time into study time and recitation time. This was done both for learning nonsense syllables and for learning meaningful continuous discourse. He found that the proportion of the time devoted to recitation for learning nonsense syllables was significantly greater than the recitation time required for learning meaningful materials. When recitation time was increased from 20% to 80% for learning nonsense syllable, the test scores were more than doubled; this same change in study-recitation time for the learning of meaningful materials had very little effect on the test scores. This result was supported by an experiment by H. A. Peterson (1944) who found that for regular reading assignments given to college students, two thirds of the time spent in reading was superior to dividing the time evenly between reading and attempted recall.

Eigen and Marguilies (50) compared overt and covert responding in memorizing lists of trigrams which varied from list to list in what they called amounts of information. One was a list of nonsense syllables, the second was a list of trigrams with higher association values, and the third, a list of common English words. Each subject was given a list and told to learn it. After a brief time it was withdrawn and the subject was presented with four of the seven trigrams. One group was required to fill in the missing items, after which his response was confirmed by presenting the entire list with the missing words underlined. This was the overt responding group. The covert group was instructed merely to think the three missing items and then consult the confirmation list. As each list was repeated the same three items were deleted.

The posttest required that all seven of the items in each list be given. The test was scored both for the practiced and unpracticed items. The overt group scored significantly higher than the covert group on both parts of the test for the nonsense syllables and the list of intermediate difficulty. But the differences were not significant for the meaningful three-letter words.

The meaningfulness of words, as measured by any technique, is related both to ease of stimulus discrimination and to response availability. It is easier to discriminate between meaningful words than between nonsense syllables. Familiar words are more available as response terms than unfamiliar ones. But these three factors appear to have one thing in common. The optimal response mode for them is overt.

5. Programming for Individual Differences in Entering Abilities.

Because of the individual differences in the entering deficiencies or abilities listed above, some students will find a program hard while others will run through it with the greatest of ease. It is possible to build into programs items that will partially correct or compensate for deficiencies such as failure to understand what the words mean, lack of the necessary ability to discriminate the key words from the context or to supply words that are missing from the student's vocabulary. But when a program is so expanded or written down to the lowest level of student entering abilities, it will inevitably become boring and tedious for the better prepared students. Susan Meyer Markle (147) has suggested that one step in the direction of individualized programming is to program the same content at three levels of difficulty, one for students who have the prerequisite entering abilities, one for those who have some but not all, and the third for those who must be taught what they need to know to learn the materials. Another plan is to program the materials for the lowest common denominator and then permit and encourage "by-passing" or "forward branching."

The ideal form of individualized instruction is, of course, the one-student-at-a-time tutorial method — or the "Hopkins-log-student" method. This method certainly requires overt responding by both the tutored and the tutor. In the language of S-R psychology this process may be described as follows: the tutor emits a stimulus, verbal or demonstrational, which stimulates the tutored first to think and then to act. The stimuli produced by the responses of the tutored elicit more thought and more response-produced stimuli from the tutor. This give and take process goes on until the tutor has moved the tutored from imperfect performances to more perfect ones. The give and take that can occur between a student and a teaching machine or a programmed textbook is very much more limited. It lacks the variety and flexibility of the tutor-tutored relationship. A computer, on the other hand, has a much greater tutorial capability than

3

a teaching machine. Herein lies the hope that the computer will be the best mechanized answer to individualized instruction.

Lewis and Pask (124) have developed a system of teaching based on the notions of cybernetics and artificially intelligent adaptive behavior. This system differs fundamentally from linear programmed instruction by the manner in which errors that obstruct learning are handled. Instead of trying to reduce error rates by prompting, the system permits errors to occur and then teaches how to avoid them. The machine is designed to detect the sources, kind, position and frequency of the occurrence of mistakes. It registers and stores a variety of error data which is used to control the level of difficulty and maintain a high level of motivation. It is described by the authors as "a device which forms a close dynamic partnership with its students in order to maximize some quantifiable measure of the student's efficiency. To this end the machine sometimes cooperates with the student to help him out of difficulty. And it sometimes competes with the student by putting new difficulties in his way." (p. 232)

The foregoing account of the contributions which overt responding may make to learning, by students who have various deficiencies in requisite entering abilities, does not cover the wider range of problems involved in individualized instruction. It represents one approach to the problem and certainly merits further investigation.

VI. SUMMARY AND OUTLOOK

A. Summary.

1. Statistical summary. During the decade 1956-1966 no less than 50 experimental comparisons between response modes have been reported. Of these 39 are comparisons between overt and covert modes, and 11 between different overt modes such as constructed versus multiple-choice responding. In 21 of the 39 no significant differences in mean posttest scores were found, in 11 the overt mode was superior, in 3 the covert was the better, and in 4 the overt was significantly better under one condition and the covert mode under another condition. (See Appendix A.) The 11 comparisons of different overt modes indicate that the differences depend to some extent on the composition of the posttest. Other comparisons have been made between overt responses such as spelling and pronouncing, writing and speaking, and again the preferred mode appears to be dependent, in some degree, on the composition of the posttests. In addition to the 50 experimental comparisons of response modes, more than 150 other experiments have been reported which contain information on how response modes interact with other conditions on which learning depends.

Statistical counts of results of experiments conducted by different investigators, on different learning tasks, using different instructional materials, presented in different formats, and using subjects selected from different populations is not very helpful for arriving at general conclusions concerning optimal response modes. The value of these experiments lies in the extent to which they reveal the conditions on which optimal response modes depend. Some of the conditions are indicated in the propositions stated below.

2. Summary of Main Conclusions.

From the extensive experimental literature cited in this paper certain tentative general propositions may be gleaned. Some are fairly well established while most are tentative and require further investigation. They are stated here rather baldly and in the order in which they are presented in the paper, and without supplying any of the qualifiers. They will serve not only to summarize the work which has already been done but also to indicate problems for further research.

In the introduction to this paper, it was pointed out that:

(1) Overt responses produce records that are indispensable for the experimental development of programs, films, ETV presentations and other instructional materials. Furthermore, overt responding is an essential requirement for all branching and computer-based programs.

They are also essential for obtaining learning curves and for revealing the processes by which students are moved from imperfect to more perfect performances. In addition to these advantages to experimenters and producers, a number of theoretical advantages to learners which have been found to be beneficial under some conditions, but not under others, were listed.

(2) In view of the fact that all forms of human learning, except that of very young children, always involve covert activities such as paying attention, observing, listening, reading, studying, and cogitating, but do not always require overt forms of behavior, the problem is not one of overt versus covert responding, as though they were independent of each other, but rather that of discovering the contributions to learning made by overt responding. When the covert processes are deliberately interfered with, the rate of learning is substantially reduced (110, 217).

(3) In programmed instruction a "thought" may be made manifest either by writing a word in a blank space, checking a multiple-choice item, speaking the response word aloud, or pressing a selector button. After a stimulus has gained control over a thought-response which can be overtly expressed in several ways, one would expect to find that the mode of expression has little or no effect on associative learning. The eleven experiments in which constructive responding has been compared with multiple-choice responding confirm this expectation.

In the section of this paper on pre-presentational conditions on which the effectiveness of overt responding may depend it was found that:

(4) Instructions given by the teacher or experimenter to students as to how they should respond are not always obeyed. Even when they are obeyed, they do not control all of the various kinds of responses that a student may make. Students who are instructed to respond overtly must read the frames first and will, no doubt, "think" of the correct response before writing it down. Those who are instructed to just "think" or to read will undoubtedly emit a variety of overt responses such as eye movements and incipient muscular responses which are not usually recorded. The contribution which all such unrecorded, and perhaps uncontrollable, overt responses make to learning is a matter beyond the limits of this paper.

(5) Pre-presentational instructions on what is to be learned has a much greater effect on optimal response modes than instructions as to how it should be learned. The contribution which overt responding can make to learning is clearly a function of what is to be learned as determined by behaviorly defined educational objectives (65, 66, 143, 235). The experimental evidence indicates that when the task is learning to discriminate the relevant from the irrelevant stimuli, the critical from the trivial, the essentials from adornments, and the key words from the context words, overt responding with KCR will greatly facilitate the

accomplishment of this task. Another form of learning to which overt responding will make substantial contributions is concept formation, which is essentially a process of stimulus discrimination on the one hand, and stimulus generalization on the other. On the response side of the S-R paradigm, the task may require the learner to add new responses to his repertory. Response-learning almost certainly requires overt responding. The same is probably true for most kinds of sequence-learning where responses are linked together in chains. But for verbal-association-learning, overt responding is of little importance provided the relevant stimuli can be discriminated and the responses are available.

Another set of conditions on which the effectiveness of overt responding appears to depend is found in the manner in which the materials to be learned are presented.

(6) When the material to be learned is programmed or semi-programmed by presenting it in the three phases of (a) observation, reading and study, (b) recitation, answering or testing, and (c) confirmation and correction, learning is more effective than it is when the second and third phases are omitted. This proposition has been found to be true in learning from instructional films (100, 128, 162, 163); from ETV presentations (29, 86, 87, 88); and from printed materials (53, 210). Overt responding in the first phase, such as reading aloud,

echoing the words of a teacher or lectures, or mimicking the actions seen, contributes little or nothing to learning.

(7) The contributions to learning which overt responding in the second phase may make depends, in part, on what happens in the first phase. If the material has been previously learned or if it is so poorly programmed that the correct answers can be given without reading it, overt responding will contribute little or nothing to posttest scores (98, 109). When the correct answers are contingent upon careful observation, close reading and understanding in the first phase, the contribution of overt responding appears to depend on how well the materials were learned during the first phase. The lower the program error rate in well programmed material, the less is the importance of overt responding (53, 80, 209).

(8) The contribution to learning of overt responding in the second phase depends also on the nature of the learning task. If the task is to copy, imitate or reproduce the stimulus materials, overt responding is much more important than when the task is to produce evidence that it was learned.

(9) The amount of material to be observed, read and studied in the first phase has a determining effect on the value of overt responding in the second. The experimental literature seems to suggest that the greater the amount of new material to be learned in the first phase, the

greater will be the contribution to posttest scores by overt answering responses (139, 140, 141, 162, 207).

(10) Schramm (185) counted ten experiments in which self-pacing was compared with fixed-rate pacing. In seven of them no significant differences were found regardless of whether the programs were presented by teaching machines, in programmed texts or television. Carpenter and Greenhill (1963) found it possible to vary the pace from 20% below to 10% above the self-paced average without having any significant effect on posttest scores. Because overt responding requires additional time, a rate of presentation that is faster than an optimal pace makes it not only more difficult but may actually interfere with learning (12, 103, 153).

(11) In laboratory experiments on paired-associate learning where learning is measured by the number of trials, or amount of time taken, to reach a criterion of one or two perfect recitations, it has been found that the shorter the exposure time of the stimulus term, the greater the number of errors per trials and the number of trials needed to reach the criterion (24, 30). In other experiments the stimulus exposure time and the responses exposure time were both held constant with the interval between the two varied (33, 34, 36). Under one condition called "prompting" the interval was .25 seconds and under another called "confirmation" it was 3.25 seconds. The shorter interval produced the fewer trials to a criterion of one perfect recitation. When the response mode was covert this superiority was significantly greater. This result supports the principle of S-R temporal contiguity.

A third set of conditions on which the contributions to learning by overt responding may depend is found in post-response events.

(12) The effects of KCR on posttest scores is no greater when the answering responses were covert than when they were overt, (162, 163).

(13) Knowledge of correct results following constructed (written) responses to frames add little or nothing to posttest scores as indicated by experiments in which KCR was given to one group and withheld from another, (95, 96, 101, 121, 143, 165, 198). When, however, the answering response mode is multiple-choice, KCR does appear to contribute significantly to posttest scores (10, 106, 225, 226).

(14) The effects on posttests and on subsequent learning of post-response stimuli, whether externally presented or response-produced, depend on the kinds of responses they elicit. Such responses are usually covert, but one experiment has been reported in which a group of young children who repeated the correct responses scored higher on the posttest than another group who was merely told that their responses were right or wrong (213). Covert responses may be rewarding and satisfying, informational, motivating, or silent rehearsal of the S-R associations.

(15) If KCR has a reinforcing or confirming effect, one would expect it to be greater when given immediately after each response than when delayed. This appears to be true (10, 59, 161). Increasing the amount of reward by money or other payments has not been found to be effective

(3, 164); but knowledge of correct answers appears to be more effective than mere knowledge of right or wrong (16, 93). Varying the schedule of reinforcement from 100% to zero, in one experiment, had no effect on posttest scores (121). Reinforcement obtained vicariously appears to have little or no effect on posttest scores (186,234).

(16) The motivating effects of KCR are almost hopelessly intertwined with other effects and with other sources of motivation. The evidence, such as it is, indicates low but positive correlations between students' attitudes toward programmed instruction and amounts learned (48,58). The same is true of audio-visual presentations (150). There is some evidence that KCR following a run of incorrect responses (massed negative reinforcement) has a depressing effect on posttest scores (160).

(17) The effects of KCR on posttest scores depends, to some extent, on the amount of useful information that is conveyed. When a student gives a wrong response, it is more beneficial to know the correct answer than to know merely that it was wrong (16). It is more helpful to give the correct response in connection with the stimulus materials (93). In some instances it is even more beneficial to give an explanation as to why it was wrong, as is done in branching programs (120). The value of post-response information depends on the uses that the learner can make of it, either on the posttest or in subsequent learning (43).

(18) Any post-response event that provides an opportunity for and encouragement to rehearse mentally the items to be learned should have a positive effect on posttest scores (23,36,120, 162).

(19) The extent to which students profit by having their mistakes corrected is indicated roughly by the percentage of items correct on the posttest which were incorrect on the programs, provided that the test items are identical with the program items. Goldbeck and Campbell (80) did this for each of three levels of program difficulty. The results were 36%, 63% and 49% suggesting that on the more difficult programs, where the error rates were high and the superior responses mode was overt, the students profited most from a knowledge of the correct answers.

A fourth set of conditions is found in individual differences in I.Q. and entering abilities.

(20) There is very little evidence to support the proposition that overt responding to programmed materials on the part of kindergarten and children in the lower grades will contribute more to learning than it will for older children. In the experiments reported by McNeil (155) and by Suppes and Ginsberg (213) indicates that overt responding is superior, but Keislar and McNeil (108) found no significant posttest difference between overt and covert responding.

(21) There is also very little experimental evidence to support the proposition that slow learners will profit more from overt responding

to programmed lessons than fast learners. The contributions which overt responding may make to posttest scores is rather independent of I.Q.'s, in the middle ranges. Whatever the dependence might be, it is masked by other important determining factors (100, 105, 122, 134, 162, 232).

(22) Although the diagnosis of entering abilities has not progressed very far (77), there is evidence which indicates that: (a) when a student lacks the ability to make the necessary stimulus discriminations between the relevant and irrelevant aspects of the materials to be learned, this deficiency may be corrected either by pre-presentation training which requires overt responding (235, 236) or by overt responding to the programmed items; (b) when the task requires responses that are not in the student's repertory, overt responding to programmed items will contribute significantly to learning (43, 230); and (c) the more a student already knows about a topic and the greater the meaningfulness of the material is to him, the less is the contribution of overt responding (50, 73).

General Propositions which are fairly well established.

(23) When learning is evaluated in terms of amounts learned per unit of time, covert responding is more efficient than covert plus overt, for the reason that overt responding always requires additional time.

(24) When instruction is conducted in the three phases of: (a) observing, listening or studying, (b) testing, answering or reciting, and

(c) confirmation and correction, overt responding in the second phase will contribute more to learning than it will in either the first or third.

(25) There is no one response mode that is optimal for all learning tasks, all media of presentation and all kinds of students. Although this proposition seems rather obvious, yet there are some advocates of both linear and branching types of programmed instruction who appear to believe that overt answering responses are essential for all forms of programmed instruction.

The main reasons why most of the above propositions remain tentative and require further investigations is that they are drawn from a diversity of fairly short experiments conducted by no less than a hundred different individuals using a wide variety of programs and other instructional materials of different lengths and presented under many different conditions to students who are far from representing scientifically selected populations, and who are posttested by many different kinds of tests, often of unknown reliability and sensitivity. Because of the diversity of exploratory character of the research thus far reported on response modes, it is not surprising that so few firm conclusions have been established.

B. Outlook

The goal of research in this area is a set of valid propositions, or rules, that are useful both to producers and consumers of instructional materials. While it is true that overt responding requires additional time, there are some educational objectives which cannot be effectively attained without having students engage in some form of overt activity. This is particularly true in the fields of mathematics, science, the fine arts, and most of the areas of vocational and professional education. The problem is to identify the educational objectives and types of learning tasks for which overt activities are essential and not a waste of time.

After a critical review of the experiments in which no significant differences were found between overt and covert modes of responding to programs, Holland (98) concluded: 'these studies undoubtedly raise enough doubts for some investigators to insure that still more research on the need for overt responses will be conducted at the expense of more fundamental variables of program design." (p. 102) The author of this statement is apparently so convinced of the odds in favor of overt responding to all programs, for all learning tasks, and by all students that further comparisons between overt and covert modes are not worthwhile (99).

Records of student responses are of indispensable value for the experimental development of instructional materials. Overt responding is essential for branching programs and for computer-based individual instruction. But the experimental work thus far reported, inadequate as it may be, clearly indicates that for linear-type programmed instruction it does not have all of the theoretical advantages flaimed for it. Five of these experiments are listed in the first section of this paper. It is true that more of the same types of experimental comparisons between overt and covert responding, as reported by most of the 39 studies, are not worthwhile. There is need, however, for more investigations of the conditions on which the effectiveness of overt responding depends (49, 132).

Listed below are some of the areas in which further research may advance knowledge toward the goal indicated above.

1. Subject-matter areas. There is need for the concentration of more experimental work on the same instructional materials or on materials selected from the same subject-matter field. What is taught and why has a determining effect on how it is best taught. The instructional materials used in the experiments reviewed in this paper range in subject-matter all the way from molecular theory taught in the third grade to programs in psychology for college students, and on to the "diagnosis of myocardial infarction" taught to nurses and laboratory technicians. The media of presentations range all the way from printed programs to instructional and training films and programmed television courses. These materials also vary in length from short programs that can be completed in a 30-minute or less to ones that require several days or even an entire semester. The results of experiments spread over such a wide range of subject-matter and media of presentation certainly do not add up to any dependable propositions of practical value to teachers on the kinds of student activities that are best suited for the teaching of any subject-matter or for

the achievement of specifiable educational objectives. Neither does it have the advantage of a random sampling of subject-matter to test how far general principles of instruction can be generalized.

In most of the experiments on response modes only one program or film has been used. There is no evidence that the results can be generalized to other programs or films. There is evidence that such generalizations would be unwarranted. For example, Hamilton and Porteus (90) replicated their experiment on long-term retention over three different programs. They found that the effects on posttest scores and on delayed retention tests of variables such as size of step, and immediate versus delayed review, were different for each of three programs. The greatest source of variance in total test scores was program differences. This result suggests the need for more experiments in which the experimental treatments are replicated over a wide range of subject-matter.

2. Individual differences. There is need for a greater concentration of research on students of different age and grade levels and particularly on those who have different levels of entering abilities. A neglected area is individual differences in previously acquired work habits, levels of aspiration and other personality traits, investigated from the standpoint of optimal response modes. It is quite likely that in working through the same program some kinds of students will gain little or nothing by overt responding while others may gain a great deal. In the experiments reviewed in this paper the participating students range all the way from

kindergarten to graduate and professional schools, and cover a very wide range of intellectual and entering abilities. It is little wonder that no firm conclusion can be drawn as to how best to move students of different ages, entering abilities and personality traits from imperfect to more perfect performances in any field of knowledge or competence. Further research in this area is especially needed for the success of computer-based instruction.

3. Analyses of the learning process. Skinner (204) suggests the need for analytical studies of learning strategies employed by different individuals in learning the same lessons. Statistical comparisons between the means of posttest scores of experimental and control groups conceal individual variations and fail to reveal the process by students are moved from imperfect to more perfect performances. He notes: "Correlations between test scores and significant differences between means tells us less about the behavior of the student in the act of learning than results obtained when the investigator can manipulate variables and assess their effects in a manner characteristic of laboratory research (p. 17)." He goes on to point out an illuminating parallel between educational and medical research. Both are aimed at improvement. But health improvement is a by-product of the changes in the specific physiological processes by which it is induced. The advancement in medical diagnosis and therapy is due,

in no small part, to the concentration of research on the physiological processes by which health is improved and maintained.

4. Learning curves. There is need for more laboratory-type research on programmed instruction. Only a very few of the experiments reported above have required each student to learn a program to a criterion of a perfect posttest score. The work by Brooks (20) on "Shaping Faster Question Answering" is one of them. Others are on paired-associate learning. One reason for making programmed learning a one trial affair is that students become bored with going over the same program twice or more. One way around this difficulty would be to prepare a ladder of programs on the same subject-matter with each program on a ladder containing fewer and fewer prompts and cues so that the final rung would be the posttest from which all prompts and cues have been vanished. From the results obtained by the use of such a ladder, a learning curve for each student could be plotted in terms of rate of program error reduction. Response modes could be covaried with ladder versus one-trial learning of the same materials (195).

5. Error analysis. Professional programmers have made extensive use of errors for the re-writing of frames and for general overall improvement. They have tabulated errors, frame by frame, and also computed the error rate for each individual student and compared the results with scores on pretest, intelligence and other measures of aptitude. But after many revisions no one has developed a program in which

the error rate made by a sample of students for whom it was intended was zero. There is some evidence that there are individual differences in what might be called "error proness" (31). This illustrates the need for going further and analyzing the kinds of errors made by different kinds of students and the places where they are made. This was done, to some extent, by Susan Meyer (161) with the errors made on her programs for teaching Latin and Greek derived prefixes to English words. In laboratory experiments on short-term memory, error analysis has proved extremely valuable for understanding the nature of this phenomenon.

6. Response analysis. If learning is by doing, as most psychologists and educators believe, and if learning depends basically on what is done in the presence of instructional materials, there is nothing of greater importance than what students actually do when they learn. Lumsdaine and May (1965) have questioned whether the critical elements in what students actually do in learning situations have been identified. When a student constructs a written answer to a frame in a linear program, he is undoubtedly doing something more than merely pushing a pencil. Likewise when a student is instructed merely to "think" the answers, he is doing more than just thinking. His thinking may be accompanied by eye movement, incipient muscular responses, knitting his brow or scratching his head. Brooks (20) points out that there are many facets of response modes which can be measured objectively, the effects of which on learning have not been

investigated. He has demonstrated that students can be trained to answer faster by simply rewarding quick answers and not rewarding slow ones even when they are correct.

7. A taxonomy of response modes. Lumsdaine (128) has suggested a fourfold classification of response modes: explicit-overt, explicit-covert, implicit-overt and implicit-covert, (p. 487). The category of explicit includes all responses which students are explicitly instructed to make to a specified category of stimuli. They are controlled by directions given prior to the presentation. For example, students may be instructed either to respond to the missing words in each frame of a program by writing in the correct word in each blank (explicit-overt) or by merely "thinking" of it (explicit-covert).

Explicit-overt responses may take any one of several forms depending on the stimulus to which the student is instructed to respond. He may be instructed to respond verbally either by speaking or writing. He may be instructed to check a multiple-choice alternative, push a selector button, draw a diagram, tie a knot, or assemble the parts of a machine. In each case the stimuli to which the student is instructed to respond must be specified, the forms of the response must be specified, and time and opportunity for the students to obey these instructions must be provided.

The category of implicit responses includes all responses that are not explicitly directed. Most implicit responses are covert, as, for example, reading, observing a demonstration, and listening to a lecture. Examples of implicit-*overt* responses are eye movement during reading, or action muscle potential recorded from the lips, or observable lip movements.

A borderline case between *explicit* and *implicit* covert responding is that of presenting frames of a program with the key words or words filled in and underlined. The student is instructed to read (implicit), but the stimuli to be noticed and responded to covertly are underlined. The underlining is, in fact, a way of explicitly defining the stimuli to which covert responses are to be made.

Gropper (82,83) points out that the terms *overt* and *covert* refer to the locus of the response, while the terms *explicit* and *implicit* refer to the specifications of the S-R relations. The terms *active* and *passive* refer to the degree of response prompting. A fully prompted response is either reading a passage or coyping it. If, on the other hand, a response is not prompted but must be anticipated from the context, as in the case of test items, it is called an *active* response . In between fully prompted and unprompted responses are a wide variety or partially prompted or "cued" responses.

In addition to these categories suggested by Lumsdaine and Gropper there are still other dimensions along which responses may be classified. For example, responses may vary from simple to complex. They also may be discrete and unitary, or an integrated complete act. Some may be instrumental or mediating and some be terminal, some may be symbolic and some operational. Each of these dimensions may be conjoined with each of the categories of explicit-implicit, overt and covert, and with each other, forming a very complicated matrix. For example, explicit-implicit, over-covert, simple-complex, mediating-terminal, symbolic-operational would constitute a 2x2x2x2x2 factorial matrix of 32 categories of response modes. In view of all the possible ways in which response modes could be classified, it is not surprising that no satisfactory taxonomy has yet been devised.

The diverse and confusing results of experiments on response variables is due, in no small part, to the lack of precise and operational definitions of them. The terms "overt" and "covert" are general and cover a variety of interrelated specifics. Some fundamental problems in psychology are here encountered. What, for example, does a student do when he thinks? What kinds of responses are involved in thinking? What functions do they perform in learning?

8. Posttests. Holland (98) has noted that one of the reasons why no significant differences were found between overt and covert responding in some experiments was that the posttests used were not sensitive enough to detect differences which may have actually existed. An ideal posttest should be not only reliable but also measure all that was learned and nothing else. The ability to give the correct answers to a test is only one form of behavior. A student may answer all questions correctly and still not fully command the material to be learned, so that he can drive in traffic with it, so to speak.

Lumsdaine (133) suggests that the ideal posttest should be composed of a sample of items drawn at random from a larger population of items, the boundaries and characteristics of which are defined by the terminal abilities which are specified by the educational objective. The reliability of such a test would be measured by the correlation between two independently drawn samples of equal length, each of which being long enough to sample adequately the entire population of items. The validity of such a posttest is the square root of its coefficient of reliability (149). If the population is composed of items requiring different abilities such as recall, recognition, transfer, and application, each of the sub-populations should, of course, be adequately sampled.

The reliability, validity and sensitivity of posttests is of crucial importance because the contributions which overt responding may make to learning are measured by the differences between means of posttest scores obtained by groups whose responses were covert only and groups whose responses were covert plus overt. Furthermore, when learning is measured by gains from pre- to posttests, it is important that the tests have a high level of reliability in order to avoid the contaminating effects of errors of measurement on gains.

The problems listed above are not the only ones that await further investigation. The vast amount of exploratory work already done opens up many new opportunities and challenges. Scientific research usually raises more questions than it answers. Research on response modes has now advanced to the stage where crude comparisons between overt and covert are no longer needed. The next stage is to discover the conditions under which overt responding will make significant contributions to learning over and above those made by the covert processes of motivation, attention, perception, cognition and comprehension. Advancement to this level has already been made by a few experiments in which response modes have been covaried with learning tasks, materials, media, pacing, confirmation, prompting and individual differences in pre-instructional variables.

Further advancement will depend largely on more long-range programs of research which are concentrated on particular aspects of general programs. This is the story of how the great advances in many fields of science have been made. A usable science of instruction and training will not be fully developed merely by the application of principles of learning derived from laboratory experiments, valuable as they may be. It will come from the development principles, rules and instructional strategies which will apply to specifiable groups of students and learning tasks.

Appendix A

I

Experiments reporting no significant differences between the means of immediate posttest scores of groups instructed to respond overtly and groups instructed either to "think" the answers or to read completed statements.

1. Alter and Silverman (4): three experiments using as 87-frame program on basic electricity with college students. No significant differences in means of posttest scores were found either when the programs were externally or selfpaced, or presented by a teaching machine or by booklets.

2. Crist, R. L. (41): used a 331-frame program on the solar system, and a 351-frame on latitude and longitude with two groups of sixth grade children. Each group responded overtly to one program and covertly to the other. No significant differences were found between response modes either on mean score of immediate posttests or on a six-weeks delayed retention test.

3. Evans (51): used a "ruleg" program of 125 frames and a less systematic one of 72 frames, both on how to construct short deductive proofs according to the rules of symbolic logic. The subjects were college students. No significant differences were found between mean scores of overt and covert responding groups either on immediate posttest or delayed retention tests.

4. Evans, Glaser, and Homme (53): using a 60-frame program on the fundamentals of music with two groups of college students (5 in each group)

found no significant differences in keen performance scores of the group instructed to write in their responses and the group not so instructed.

5. Feldhusen and Birt (57): used several formats of linear programs on teaching machines and in booklets. One format consisted of complete frames to be read, and another required overt answering. No significant differences in mean posttest scores were found either for this or for any of the other comparisons which suggest that an insensitive posttest was used.

6. Groper and Lumsdaine (86): found that writing responses to a programed television course on body chemistry resulted in no better posttest mean score than reading the items with the blanks filled in.

7. Gropper and Lumsdaine (88): made another comparison between active and passive responding to a non-programed televised lesson on Newton's laws and found no significant difference on mean posttest scores of subjects who were and were not encouraged to make active responses.

8. Hartman, Morrison, and Carlson (91): found no significant differences in immediate posttest scores between groups instructed to write in answers, and groups instructed to read complete frames of a program designed to teach IBM operations.

9. Hughes (102): using a 719-frame linear program on basic computer knowledge with nine classes of IBM trainees found no significant posttest differences between groups instructed to write in answers and groups not so instructed.

10. Kanner and Sulzer (107): found no significant differences between means of posttest scores of groups who responded either orally, in writing, or by "thinking" answers in learning the army phonetic alphabet.

11. Kaess and Zeaman (106): found that in learning psychological vocabulary by the use of a simulated form of the Pressey punchboard, reading the correct multiple choice alternatives was about as effective as punching them.

12. Keislar and McNeil (108): using a 432-frame linear program on physical science with children in the primary grades, found no significant posttest differences between mean scores of groups instructed to respond overt-orally and groups who responded covertly.

13. Krumboltz and Weisman (121): using a 177-frame linear program on educational testing with college students found no significant difference between overt and covert responding groups on the immediate posttest, but on a two weeks delayed retention test the "overts" had a significantly higher mean score than the "coverts".

14. Lambert, Miller and Wiley (122): covaried overt and covert modes with levels of intelligence using an 864-frame linear program on sets, relations and functions with 522 ninth grade students and found no significant differences between response modes on the immediate posttest.

15. Maccoby, Michael and Levine (163): covaried overt and covert modes with confirmation versus no confirmation using two instructional

films with Air Force trainees and found no significant difference on immediate posttest between "overts" and "coverts" but a very significant difference in favor of immediate confirmation.

16. Kendler, Cook and Kendler (110): found that pre-film instructions either to participation answers to semiprogramed film or to "think" them was better than no instructions, but no significant difference on immediate posttest mean scores resulted from instructions to write or to think.

17. Roe (177): using a 230-linear frame program of elementary probability, found that engineering students who were instructed to respond overtly did no better on the immediate posttest than those who were not so instructed.

18. Shettel and Lindley (192): found that in teaching the army phonetic alphabet to college students the covert study of flash cards was about as effective as overt responding to a 90-frame constructed-response program.

19. Stolurow and Walker (209): Using a program on descriptive statistics found that college students who were instructed to "think" the answers had a mean posttest score about as high as those who were instructed to write in the answers.

20. Tobias and Weiner (219): found no significant differences either on immediate posttests or delayed retention test between groups who were instructed to write, think, or read completed frames of a review program.

21. Wynn and McKeegan (238): compared overt responding to a program designed to teach Federal Relations to Education to graduate law students with covert responding to the same program and with discussion of cases. On a multiple-choice achievement test there was no significant difference between the "overts" and "coverts" and no difference on a four to five weeks delayed retention test.

Appendix A

II

Experiments in Which Overt Responding Was Found Superior to Covert.

1. Csanyi, Glaser and Reynolds (42): compared oral with thinking responses to a 480-frame linear program on teaching the pronunciation of 12 phonetic symbols and found that oral responding yielded a posttest mean score significantly higher than that yielded by "thinking" the answers.

2. Cummings and Goldstein(43): covaried the response modes of writing versus thinking the answers to 119-frame linear program on the diagnosis of myocardial infarction with verbal and pictorial stimulus materials and found that student nurses and technicians who wrote in the answers scored higher on the verbal and pictorial posttest tests and delayed retention test than groups who were instructed to "think" the answers.

3. Gropper and Lumsdaine (87): found that when a television lesson on the movies was programed in a manner similar to teaching machine programs, students who participated actively (i.e., responded overtly) performed significantly better on the posttest than those who merely watched the program.

4. Holland (95): using 22 sections of the Holland-Skinner program on psychology, compared the usual written mode of responding plus

confirmation with reading a version of the same program with the answers filled in and found that the reading group made significantly more errors on the posttest.

5. Hovland, Lumsdaine and Sheffield (100): in their classic experiment on teaching the army phonetic alphabet to soldiers found that groups who were instructed to participate actively (i.e.,to responding by answering aloud and in concert) during the review sessions performed significantly better on the posttest than groups who were instructed to remain silent.

6. Lumsdaine and Gladstone (134): repeated the Hovland, Lumsdaine and Sheffield experiment, using a more simplified version of the same instructional materials and found that active participation during review sessions produced better learning than silent listening and observing both for subjects of high and low intellectual abilities.

7. McNeil (155): using a 700-frame linear program on reading skills, found that kindergarten children who were instructed to pronounce each word as it was presented learned significantly better than groups who were instructed just to look at each word.

8. Mechanic and D'Andrea (158): found that in learning which of the three letters of a trigram had been selected to be "correct", groups who were instructed to either to pronounce or to spell each trigram aloud did significantly better on the posttest than groups instructed to pronounce or spell silently.

9. Willians (230): using some of the sections of the Holland-Skinner program on psychology for purposes of review found that overt-constructed responding resulted in significantly better posttest scores than reading completed statements. This was particularly true of frames that required the learning of new technical terms.

10. Wulff and Emeson (235): also found that overt responding during training to discriminate between drawings of electric circuits resulted in better learning of the names attached to each drawing than mere studying the drawings for 20 minutes.

11. Wulff and Kraeling (236): found that overt responding in pre-film stimulus discrimination training resulted in better learning from a film than covert responding during the pre-film training period.

Appendix A

III

Experiments in Which the Covert Mode Was Found Superior to the Overt.

1. Cook and Spitzer (36): compared overt constructed responding with covert under both prompting and confirmation procedures and found that the fastest learning was obtained under covert-prompting and the slowest under overt-confirmation.

2. Sidowski, and others (193): using a program of 15 Russian-English paired vocabulary items, covaried overt and covert response modes with prompting and confirmation procedures and found that for groups of college students covert responding was the more effective especially under the prompting condition.

3. Silberman, Malarngo, and Coulson (198): using a 61-frame linear multiple-choice program on logic with each frame presented on a card, found that high school students who were instructed to study the cards scored significantly higher on the posttest than those who were instructed to respond overtly to the items.

Appendix A

IV

Experiments in Which Covert Responding Was Found Superior Under One Condition and Overt Superior Under Another Condition.

1. Eigen and Margulies (50): covaried overt and covert response modes with levels of difficulty inherent in list of trigrams. Difficulty defined as amounts of information conveyed by each item. Nonsense syllables considered to be more difficult than common three-letter words. Overt responding was the more effective for learning the most difficult, and item of intermediate difficulty. But covert responding was better for the least difficult - i.e., the items conveying the most information.

2. Goldbeck and Campbell (80): covaried response modes of writing, thinking, reading with three programs of different levels of difficulty. At the easiest level "thinking" and reading resulted in significantly higher posttest scores than writing in the answers; at the intermediate level the writing mode was superior to the thinking and reading modes. At the most difficult levels the differences between the three levels was slight.

3. McGuire (153): covaried overt and covert responding with rates of presentation of slides designed to teach the names of nine mechanical parts. The overt mode was superior at the slow rate; but the covert was better at the faster rate.

4. Wittrock (232): using a program designed to teach the relations of molecular action to the phenomena of evaporation and condensation to first and second grade children, found no significant difference between overt and covert responding on the part of children whose I.Q.'s were above the median, but the overt mode was significantly better for those with I.Q.'s below the median.

REFERENCES

1. Allen, W. H. Research on Film Use: Student participation. AV Communication Rev., 1957, 5, 423-450.

2. Alter, Millicent. Retention in programmed instruction. New York. Center for Programmed Instruction, 1962. (offset)

3. Alter, Millicent; Eigen, Lewis and King, Shirley. The effectiveness of confirmation plus trinket reinforcers in young children. New York: Center for Programmed Instruction, Inc., 1962, 22 pp. (offset)

4. Alter, Millicent and Silverman, Robert. The response in programmed instruction. Journ. Programmed Instruction, 1, 1 (1962), 55-78.

5. Angell, David and Lumsdaine, Arthur A. A study of subject-controlled partial cueing in paired associate learning. San Mateo, California: American Institutes for Research, 1961. (b) 13 pp. (mimeo.)

6. Angell, David and Lumsdaine, Arthur A. The effects of prompting trials and partial-correction procedures on learning by anticipation. San Mateo, California: American Institutes for Research, 1961, (a) 47 pp. (mimeo.)

7. Angell, David and Lumsdaine, Arthur A. Retention of material presented by autoinstructional programs which vanish and which do not vanish verbal cues. Palo Alto, California: American Institutes for Research, 1962. 13 pp. (mimeo.)

8. Angell, David and Lumsdaine, Arthur A. Prompted plus unprompted trials versus prompted trials alone in paired-associate learning. Pittsburgh, Pa.: American Institutes for Research, 1960. 19 pp. (mimeo.)

9. Angell, David and Lumsdaine, Arthur A. Research on Cueing Factors Related to Programmed Instruction. Pittsburgh, Pa. American Institutes for Research, Report C-14-9/62. 1962.

10. Angell, G. W. The effect of immediate knowledge of quiz results and final examination scores in freshman chemistry. Journ. Educ. Research, 42 (1949), 391-94.

11. Annet, John. The role of knowledge of results in learning: A survey. In: Educational Technology (Dececco, J. P., Ed., Holt, Rinehart and Winston, 1964, 479 pp.

12. Ash, P. and Carlton, B. J. The Value of Note-Taking During Film Learning (Pennsylvania State Univ. Instruct. Film Res. Program, U. S. Naval Train. Device Cent., ONR. Tech. Rept. No. SDC 269-7-21, Port Washington, N. Y., 10 pp., 1951).

13. Beane, D. G. A comparison of linear and branching techniques of programmed instruction in plane geometry. Urbana, Ill., University of Ill. Tech. Rept. No. 1 1962.

14. Black, H. B. The effectiveness of observing response in programmed pictorial instruction. Paper read at 1965 Convention of DAVI, mimeo, 8 pp.

15. Briggs, Leslie J. Self-pacing versus automatic pacing in paired-associate learning. In Lumsdaine, (Ed.) Student Response in Programmed Instruction, 1961, 399-409. (See Ref - 128)

16. Briggs, Leslie J. Prompting and confirmation conditions for three learning tasks employing the subject-matter trainer. In: Lumsdaine (Ed.), Student Response in Programmed Instruction, 1961, 375-387. (See Ref. 128)

17. Briggs, L. J. and Angell, D. Programmed instruction in science and mathematics. Rev. of Educational Research 34. 1964, 354-372.

18. Briggs, L. J. and Hamilton, N. R. Meaningful learning and retention; practice and feedback variables. Rev. of Educational Research 34 (5) 1964 545-558.

19. Brooks, Lloyd O. Response latency in programmed learning: Latency related to error rate. Ph.D. dissertation, University of Houston, 1961.

20. Brooks, L. O. Shaping faster question answering. Palo Alto, Calif. American Institutes for Research. Tech. Rept. No. 10. to Office of Naval Research 1965, 65 pp. (mimeo.)

21. Brown, G. H. Factors influencing incidental learning. J. Exp. Psychol. 47 1954 163-169.

22. Bryan, G. L. and Rigney, J. W. *An evaluation of a method for shipboard training in operations knowledge.* Los Angeles: Department of Psychology, University of Southern California, 1956.

23. Bryan, Glenn, L., Rigney, Joseph W. and Van Horne, Charles. *An evaluation of three types of information for supplementing knowledge of results in a training technique.* Los Anegeles: Electronics Personnel Research Group, Department of Psychology, University of Southern California, 1957, 23 pp. (mimeo.)

24. Bugelski, R. R. *Presentation time, total time and mediation in paired-associate learning.* J. Exp. Psychol. *63* 1962, 409-412.

25. Burton, Benjamin B. and Goldbeck, Robert A. *The effect of response characteristics and multiple-choice alternatives on learning during programmed instruction.* San Mateo, California: American Institutes for Research, 1962, 16 pp. (offset).

26. Campbell, Vincent, N. Adjusting self-instruction programs to individual differences: studies of cueing, responding, and bypassing. San Mateo, California: American Institute of Research, 1961, 56 pp. (offset).

27. Campeau, Peggie L. Level of Anxiety and Presence or Absence of Feedback in Programmed Instruction. Palo Alto, Calif.; American Institutes for Research, 1965 35 pp. (mimeo.)

28. Campeau, Peggie L. Selective review of literature on audio-visual media of instruction. (In Briggs, L. J. and others, Instructional Media: A Procedure for the Design of multi-media Instruction, A Critical Review of Research. Palo Alto, Calif. American Institutes for Research 1965)

29. Carpenter, C. R. and Greenhill, L. P. et al. Comparative Research on Methods and Media for Presenting Programmed Courses in Mathematics and English. University Park, Pa.: University Division of Instructional Services, The Pennsylvania State University, 1963, 74 pp.

30. Carroll, J. B. and Burke, M. L. Parameters in paired-associate verbal learning: List length, meaningfulness, rate of presentation and ability. J. of exp. Psychol., 1965, 69, 543-553.

31. Cook, Donald, Studying the performance of a program. In deGrazia and Sohn (eds) Programs, Teachers, and Machines. New York, Bantam Books, 1964

32. Cook, Donald, and Mechner, F. Fundamentals of Programmed Instruction. (In Margulies and Eigen (eds) Applied Programmed Instruction New York, John Wiley & Sons, 1962.)

33. Cook, John Oliver. Supplementary report: processes underlying learning a single paired-associate item. J. exp. Psychol., 56, 1958, 455 pp.

34. Cook, John Oliver. Response analysis in paired-associate learning experiments. In Lumsdaine (Ed.), Student Response in Programmed Instruction, 1961, Washington, D. C. National Academy of Science 367-374

35. Cook, John Oliver and Kendler, Tracy S. A theoretical model to explain some paired-associate learning data. In Symposium on Air Force human engineering, personnel and training research, Glen Finch and F. Cameron (Eds.). (National Research Council Publications, 455 Washington, D. C.: NRC, 1958, 90-98 pp.)

36. Cook, John Oliver and Spitzer, Morton E. Supplementary report: Prompting versus confirmation in paired-associate learning. *J.exp. Psychol.*, *59*, 1960, 275-76 pp.

37. Coulson, J. E. (ed.) *Programmed Learning and Computer-Based Instruction.* New York, John Wiley & Sons. 1962.

38. Coulson, E. and Silberman, F. Effects of three variables in a teaching machine. *J. educ. Psychol.*, *51*, 1960, 135-43 pp.

39. Coulson, J. E. and others. Effects of branching in a computer autoinstructional device. *J. Applied Psychol.* *46*, 1962, 389-92

40. Corbin, M. T. and McIntyre, C. J. *Development and application of a new method to test the relative effectiveness of specific visual production techniques for television.* Urbana, Ill. Univ. of Ill. 1961

41. Crist, R. L. Overt versus covert responding and retention by sixth-grade students. *Jour. Educ. Psychol.* *57* 1966, 99-101.

42. Csanyi, A. P. Glaser R., and Reynolds, J. H. Programming methods and Response mode in a visual-oral response task.

(In _Investigation of learning variables in programmed instruction._) Pittsburgh, Pa. Programmed Learning Laboratories Univ. of Pittsburgh, 1962.

43. Cummings, A. and Goldstein L. S. The effects of overt and covert responding on two kinds of learning tasks. In: _Educational Technology_, DeCecco, J. P. (ed.) New York, Holt, Rinehart and Winston, 1964

44. Davidson, R. E. Mediation and Ability in paired-associate learning. _J. educ. Psychol._, 1964, _56_, 352-356

45. DeCecco, John P. (Ed.) _Educational Technology: Readings in Programmed Instruction._ New York: Holt, Rinehart & Winston, 1964

46. Deese, James. Behavioral Effects of Instruction to Learn. In Melton, A. W. (ed.) _Categories of Human Learning._ New York. Academic Press 1964

47. Eigen, L. D. A Comparison of three modes of presenting a programmed instructional sequence. _J. educ. Psychol._, _55_, 1962, 453-460

48. Eigen, L. D. High school student reaction to programmed instruction. _Phi Delta Kappan_ _44_ 1963, 282-85.

49. Eigen, L. D. Problems of Research in programmed instruction. _AV Comm. Rev. 13 (1) 1965_

50. Eigen, Lewis D. and Margulies, Stuart. Response characteristics as a function of information level. New York: Center for Programmed Instruction, n.d., 1963, (offset).

51. Evans, James L. An investigation of "teaching machine" variables using learning programs in symbolic logic. Doctor's dissertation, University of Pittsburgh, 1960.

52. Evans, James L. Programming in mathematics and logic. (In Glaser, R. (ed.) Teaching Machines and Programmed Learning, II Data and Direction. pp. 371-440) (see ref. - 77)

53. Evans, James L., Glaser, Robert and Homme, Lloyd E. An investigation of variation in the properties of verbal learning sequences of the "teaching machine" type. In: Lumsdaine and Glaser (Eds.), Teaching Machines and Programmed Learning: A source book (Washington, D. C., Department of Audio-Visual Instruction, National Education Association, 1960, 446-51 pp.)

54. Evans, James L., Homme, L. E. and Glaser, R. The Ruleg System for the construction of programmed verbal learning sequences. J. Ed. Res., 1962, 55, 513-518 pp.

55. Faison, E. W. J., Rose, N. and Pudell, J. E. The effects of rest pauses during training film instructions. Air Force Personnel Training Center, Trussy Aids Research Laboratory, Tarl-LN-55-44, 1950 (a)

56. Faison, E.. W. J., Rose, N. and Pudell, J. E. A technique for measuring observable audience reactions to training films. TARL-LN-55-46, 1955 (b)

57. Feldhusen, John F. and Birt, Andrew. A study of nine methods of presentation of programmed learning material. J. Educ. Research, 55, 1962, 461-66 pp.

58. Feldhusen, J. F. and Eigen L. D. Interrelationships Among Attitude, Achievement, Reading, Intelligence, and Transfer Variables in Programmed Instruction. Paper presented to the Midwestern Psychological Association, 1963, 23 p.

59. Flemming, M. L. Feedback to programmers, participation by learners, and feedback to learners. Bloomington, Ind., Ind. University School of Education, Divisions of Ed. Media (1963)

60. Follettie, J. F. Effects of Training Response Mode, Test Form, and Measure on Acquisition of Semi-Ordered Factual Material. (Army Infantry Human Res. Unit, Fort Benning, Ga. 1961 62 pp.)

61. Fry, Edward B. A study of teaching machine response modes. In: Lumsdaine and Glaser, (Ed.). Teaching Machines and Programmed Learning: A Source Book. Washington, D. C.: Department of Audio-Visual Instruction, National Educational Association, 1960, 469-74 pp.

62. Frye, Charles H. Group vs. individual pacing in programmed instruction. Oregon State System of Higher Education, 1963

63. Gage, N, L. (ed.) *Handbook of Research on Teaching.* Chicago, Rand McNally & Co. 1963

64. Gagne, R. M. The acquisition of knowledge. Pacyhol. Rev. *69* 1962 355-365.

65. Gagne, R. M. *The Conditions of Learning.* New York, Holt, Rinehart and Winstin 1965

66. Gagne, R. M. Analysis of instructional objectives. In Glaser, R., (ed.) *Teaching Machines and Programmed Learning II.* 21-65.

67. Gagne, R. M. and Bolles, R. C. A Review of the factors in learning efficiency. In Galanter, E. (ed.) *Automatic Teaching: The State of the Art.* 13-54. New York: Wiley and Sons 1959

68. Gagne, R. M. and Bassler, O. Study of retention of some topics of elementary nonmetric geometry. *Jour. Eudc. Psychol.*, *54*, 1963, 123-131.

69. Gagne, R. M. and Dick, W. Learning measures in self-instructional solving equations. *Psychol. Reports,* *10*, 1962 131-146.

70. Gagne, R. M. and Smith E. C. Jr. A study of the effects of verbalization on problem solving. *Jour. Exper. Psychol.* *62*, 1962, -12-18

71. Gagne, R. M. and Paradise, N. E. Abilities and learning sets in knowledge acquisition. *Psychol. Monographs Vol. 75 no. 14, Whole No 518.* 1961

72. Galanter, E. Automatic Teaching: The State of the Art. New York: John Wiley and Sons, 1959

73. Gates, A. I. Recitation As a Factor in Memorizing. New York: Columbia University, Archives of Psychology 7 1917

74. Gilbert, T. F. On the relevancy of laboratory investigation of learning to self-instructional programming. (In Lumsdaine, A.A. and Glaser, R. (eds.) Teaching Machines and Programmed Learning. Washington, D. C. DAVI, NEA. 1960)

75. Gilbert, T. F. Mathetics: The Technology of Education. Jour. Mathetics I, 1962

76. Gilbert, T. F. Programming Blueprinted Contemp. Psychol. 9 1964 378-80

77. Glaser, R. (ed.) Teaching Machines and Programmed Learning II; Date and Directions. Washington, D.C. DAVI, NEA. 1965

78. Glaser, R. and Reynolds, J. H. Investigation of Learning Variables in Programmed Instruction, Univ. Pittsburgh, Programmed Learning Lab., Pittsburgh, Pa., 182 pp., October 1962

79. Glaser, R. and Taber, J. I. Investigations of the Characteristics of Programmed Learning Sequences, Univ. Pittsburgh, Programmed Learning Lab., Pittsburgh, Pa., 186 pp., 1961.

80. Goldbeck, R. A. and Campbell V. N. The effects of response mode and response difficulty on programmed learning. J. Educ. Psychol. 53 1962 110-118

81. Goldstein, L. S. and Gotkin, G. A review of research: Teaching machines versus programmed textbooks as presentation modes. J. Programmed Instruction, 1 (1) 1962, 29-31

82. Gropper, G. L. A Behavioral Analysis of the Role of Visuals In Instruction. Pittsburgh, Pa. Metropolitan Pittsburgh Educational Television Station and the American Institutes for Research 1963

83. Gropper, G. L. Controlling Student Response During Visual Presentations. Metropolitan Pittsburgh Educational Television Station and American Institutes for Research. 1965 Pittsburgh, Pa.

84. Gropper, G. L. and Kress, G. C. Individual Instruction through pacing procedures. AV Comm. Rev. 13 (2) 1965, 165-182.

85. Gropper, G. L. Learning from visuals: Some behavioral considerations. AV Comm. Rev. 14 (1) 1966

86. Gropper, George L. and Lumsdaine, Arthur A. An experimental comparison of a conventional TV lesson with a programmed TV lesson requiring active student response. Studies in Televised Instruction, Report No. 2. Pittsburgh, Pa.: Metropolitan Pittsburgh Educational Television Stations WQED-WQEX, and American Institute for Research, 1961 (a) 21 pp. (offset)

87. Gropper, George L. and Lumsdaine, Arthur A. An experimental evaluation of the contribution of sequencing, pre-testing and active student response to the effectiveness of "programmed" TV instruction.

Studies in Televised Instruction, Report No. 3 Pittsburgh, Pa., Metropolitan Pittsburgh Educational Television Stations WQED-WQEX, and the American Institute for Research, 1961 (b), 19 pp. (offset)

88. Gropper, Geroge L. and Lumsdaine, Arthur A. An investigation of the role of selected variables in programmed TV instruction. Studies in Televised Instruction, Report No. 4 Pittsburgh: Metropolitan Pittsburgh Educational Television Stations WQED-WQEX and American Institute for Research, 1961, 29 pp. (offset)

89. Hamilton, N. R., Briggs, L. J., and Porteus, B. D. Improvement of retention through review. Palo Alto, Calif. Amer. Inst. for Research

90. Hamilton, N. R. and Porteus, B. D. Instruction from small step versus large step programs with immediate versus delayed review procedures for students of different ability levels. Palo Alto, Calif., Amer. Instit. for Research, 1965, 75 p. mimeo.

91. Hartman T. F., Morrison, B. A. and Carlson, M. E. Active responding in programmed learning materials. J. Applied Psychol. 47 343-47, 1963

92. Heimer, R. T. Comparison of External pacing and self-pacing of programmed instruction in mathematics using different methods of

presentation. (In Carpenter, C. R. and Greenhill, L. P. (eds) Comparative Research on Methods and Media for Presenting Programmed Courses in Mathematics and English. University Park, Pa., Penn. State University, 1963

93. Hirsch, R. S. The Effects of Knowledge of Test Results on Learning of Meaningful Material (Pennsylvania State Univ. Instruct. Film Res. Program, U. S. Naval Train. Device Cent., ONR, Tech. Rept. No. SDC 269-7-30, Port Washington, N. Y., 27 pp., September 1952

94. Hoban, C. F. Jr., and Van Ormer, E. B. Instructional Film Research, 1918-1950. (Pennsylvania State Univ. Instruct. Film Res. Program, U. S. Naval Train. Device Cent., ONR. Tech. Rept. No. SDC 269-7-19, Port Washington, N. Y., 179 pp., 1950)

95. Holland, J G. Design and Use of a Teaching Machine Program (Paper read at symposium on automated teaching methods in undergraduate instruction and educational research, Am. Psychol. Assoc., Chicago, September, 1960) (Mimeo., unpubl.)

96. Holland, J G. Teaching machines: an application of principles from the laboratory. In: Lumsdaine, A.A. and Glaser, R. (eds.) Teaching Machines and Programmed Learning, Washington, D.C. DAVI, NEA 1960

97. Holland, J. G. and Porter, D. The influence of repetition of incorrectly answered items in a teaching-machine program.

J. Exptl. Anal. of Behav., 4, 305-7, 1961

98. Holland, J. G. Research on Programming Variables. (In Galser, R. (ed.) Teaching Machines and Programmed Learning II. Washington, D. C., DAVI, NEA 1965)

99. Holland, J. G. and Skinner, B. F. (eds.) An Analysis of the Behavioral Processes Involved in Self-Instruction with Teaching Machines. Project No. 191, NDEA, Title VII 1965)

100. Hovland, C. I., Lumsdaine, A. A. and Sheffield, F. D. Experiments in Mass Communication. Princeton, N. J. Princeton Univ. Press 1949

101. Hough, John B. and Revsin, Bernard. Programmed instruction at the college level: a study of several factors influencing learning. Phi Delta Kappa, 44, 6, 1963, 286-91 pp.

102. Hughes, J. L. Effect of changes in programmed text format and reduction of classroom time on the achievement and attitude of industrial trainees. J. Programmed Instruction 1 (1) 1962, 43-54

103. Jaspen, N. Effects on training of experimental film variables, Study II: Verbalization, "How it works," Nomenclature, Audience Participation, and Succint Treatment (Pennsylvania State Univ. Instruct., Film Res. Program, U. S. Naval Train. Device Cent., ONR, Tech. Rept. No. SDC 269-11, Port Washington, N. Y., 25 pp. March 1950)

104. Jenkins, J. G. Instruction as a factor in incidental learning. *Amer. Jour. of Psych.* *45*, 1953 471-477

105. Jensen, A. Learning ability in retarded, average and gifted children. Merrill-Palmer Quarterly of Behavior and Development. *9* (2) 1963, 124-140.

106. Kaess, Walter and Zeaman, David. Positive and negative knowledge of results on a Pressey-type punchboard. *Journ. Exp. Psychol.,* *60*, 1960, 12-17 pp.

107. Kanner, Joseph H. and Sulzer, Richard L. Overt and covert rehearsal of fifty percent vs. one hundred percent of the material in film learning. In: Lumsdaine, (Ed.), 1961, 427-441.

108. Keislar, E. R. and McNeil, J. D. A comparison of two response modes in an auto-instructional program with children in the primary grades. *J. Educ. Psychol.,* 1962, *53*, 127-131 pp.

109. Kemp, F. D. and Holland, J. G. Blackout ratio and overt responses in programmed instruction: Resolution of disparate results. J. Educ. Psychol. *59* 1966, 109-114

110. Kendler, T. S., Kendler, H. H., and Cook, J. O. Effect of opportunity and instructions to practice during a training film on initial recall and retention. AFPTRC Training Aids Research Laboratory. Staff Memo 29. 1954

111. Kimble, Gregory A. and Wulff, J. Jepson. The effectiveness of instruction in reading a scale as influenced by the relative amounts

of demonstration and problem solving practice. In: Lumsdaine, (Ed.) Student Response in Programmed Instruction 1961, 227-239 pp.

112. Kimble, Gregory A. and Wulff, J. Jepson. "Response guidance" as a factor in the value of student participation in film instruction. In: Lumsdaine (Ed.), 1961, 217-226 pp.

'113. Klaus, D. J. An Analysis of programming techniques. (In Glaser, R. (ed.) Teaching Machines and Programmed Learning, II. Washington, D. C. DAVI, NEA 1965

114. Knight, H. R. and Sassenrath, J. M. Relation of achievement motivation and test anxiety to performance in programmed instruction. J. Educ. Psychol. 57 (1), 1966 14-17

115. Kopstein, Felix F. and Roshal, Sol. M. Verbal learning efficiency as influenced by the manipulation of representational response processes: pictorial-verbal and temporal contiguity factors. In: Lumsdaine, (Ed.), Student Response in Programmed Instruction, 1961, 335-50 pp.

116. Kormondy, E. J. and Van Etta, E. L. Experimental self-instruction in general biology. Ohio Jour of Science 4 1962 4-10

117. Kress, G. C. Jr. and Gropper, G. L. A comparison of two strategies for individualizing fixed-paced programmed instruction. Proceedings of the 73rd Annual Convention of the APA. APA, Washington, D. C., 1965

118. Krumboltz, John D. and Weisman, Ronald G. The effect of overt vs. covert responding to programmed instruction on immediate and delayed retention. Jour. Educ. Psychol., 53, a 2, 1962, 89-92 pp.

119. Krumboltz, John D. The nature and importance of the required response in programmed instruction. Paper presented to American Educational Research Association, Chicago, February, 1963. 11 pp. (mimeo.)

120. Krumboltz, J. D. and Bonawvitz, Barbara. The effects of receiving the confirming response in context in programmed material. J. Educ. Res., 1962, 55, 472-475 pp.

121. Krumboltz, John D. and Weisman, Ronald G. The effect of intermittent confirmation in programmed instruction. Journ. Educ. Psychol., b 53, 6, 1962, 250-53 pp.

122. Lambert, Philip, Miller, Donald M. and Wiley, David E. Experimental folklore and experimentation: The study of programmed learning in the Wauwatosa Public Schools. Journ. Educ. Research, 55, 1962, 485-94

123. Lane, Harlan, Programmed Learning of a second language. (in Glaser, R. (ed.) Teaching Machines and Programmed Learning, II. Washington, D. C. DABI, NEA 1965)

124. Lewis, B. N. and Pask G. Theory and practice of adaptive teaching systems. (In Glaser R. (ed.) Teaching Machines and Programmed

Learning, II. Washington, D. C. DAVI, NEA, 1965

125. Little, J. K. Results of use of machines for testing and drill upon learning in educational psychology. J. Exp. Educ. 3 1934, 45-49.

126. Lumsdaine, A. A. Partial and more complete automation of teaching in group and individual learning situations. In: Automatic Teaching: The State of the Art, 147-66 (Galanter, E. H., Ed., Wiley, New York, 198 pp. 1959)

127. Lumsdaine, A. A. Some issues concerning devices and programs for automated learning. In: Lumsdaine and Glaser (Eds.) Teaching Machines and Programmed Learning, 517-34 Natl. Educ. Assoc., Washington, D. C. 1960

128. Lumsdaine, A. A., Ed. Student Response in Programmed Instruction (Natl. Acad. Sci., Natl. Res. Council, Washington, 555 pp., 1961

129. Lumsdaine, A. A. Some theoretical and practical problems in programmed instruction. In: Programmed Learning and Computer-Based Instruction, 134-51 (Coulson, J. E., (Ed.) Wiley, New York, 219 pp., 1962)

130. Lumsdaine, A. A. Instruments and media of instruction. In: Handbook of Research on Teaching, 583-682 (Gage, N. L., Ed., Rand-McNally, Chicago, 1219 pp., 1963^{a}

131. Lumsdaine, A. A. Some problems in assessing instructional programs, In: *Prospectives in Programming*, 228-62 (Filep T. T., Ed., Macmillan, New York, 324 pp., 1963)

132. Lumsdaine, A. A. Educational technology, programmed learning, and instructional science. In: (Hilgard, E. R., Ed.,) *Theories of Learning and Instruction: Sixty third Yearbook of the National Society for the Study of Education*, 371-401 Univ. Chicago Press, Chicago, 430 pp., 1964

133. Lumsdaine, A. A. Assessing the effectiveness of instructional programs. In Glaser (ed.) *Teaching Machines and Programmed Learning, II.* P. 267-320 (See Ref. 77)

134. Lumsdaine, A. A., and Gladstone, A. I. Overt practice and audio-visual embellishments. In *Learning from Films*, 58-71 (Lumsdaine, A. A. and May, M. A., (Eds.), Yale Univ. Press, New Haven, Conn., 357 pp., 1958)

135. Lumsdaine, A. A. and May, M. A. Mass Communication and Educational Media. Annual Review of Psychology 1965 475-534

136. Lumsdaine, A. A., and Glaser, R., Eds. *Teaching Machines and Programmed Learning* (Natl. Educ. Assoc., Washington, 724 pp., 1960).

137. Lumsdaine, A. A., Sulzer, R. L. and Kopstein, F. F. The effect of animation cues and repetition of examples on learning from an instructional film. In Student Response in Programmed Instruction, 241-69 (Lumsdaine, A. A., Ed., Natl. Acad. Sci., Natl. Res. Council, Washington, 555 pp., 1961).

138. Lubin, Irwin. Source differences in effectiveness among controllers of verbal reinforcement. Proceedings of 73rd Annual Convention, APA. APA Washington, Dec., 1965.

139. Maccoby, Nathan; Michael, Donald N. and Levine, Seymour. Further studies of student participation procedures in film instruction: review and preview covert practice, and motivational interactions. In Lumsdaine, (Ed.), 1961, 295-325 pp. (See Ref. 128)

140. Maccoby, N. and Sheffield, F. D. Theory and experimental research on the teaching of complex sequential procedures by alternate demonstration and practice. In Symposium on Air Force Human Engineering, Personnel, and Training Research, 99-107 (Finch, G., and Cameron, F. (Eds.), Natl. Acad. Sci. Natl. Research Council, Washington, 216 pp., 1958)

141. Maccoby, N. and Sheffield, F. D. Combining practice with demonstration in teaching complex sequences: summary and interpretation. In Student Response in Programmed Instruction, 1961, 77-85.

142. MacDonald, F. J. and Allen, D. W. An investigation of presentation response and correction factors in programmed instruction. *J. Educ. Res.*, 1962, *55*, 502-507 pp.

143. Mager, R. F. *Preparing Objectives For Programmed Instruction.* San Francisco, Calif. Fearon Press 1962

144. Margolius, Garry J., Sheffield, F. D., Maccoby, Nathan. Timing of demonstrations and overt practice as a function of task organization. (In Lumsdaine (Ed.), 1961, 101-105 pp.)

145. Margolius, Garry J., Sheffield, Fred D., and Maccoby, Nathan. Repetitive versus consecutive demonstration and practice in the learning of a serial mechanical-assembly task. In Lumsdaine, (Ed.)., 1961, 87-100 pp.

146. Margolius, Garry J., and Sheffield, Fred D. Optimum methods of combining practice with filmed demonstration in teaching complex response sequences: serial learning of a mechanical assembly task. In Lumsdaine, (ed.), 1961, 33-53 pp.

147. Markle, Susan Meyer. Individualizing programmed instruction: The programmer's part. In, de Grazia, A. and Sohn, D. A. (eds.) *Programs, Teachers, and Machines,* New York. Bantam Books 1964.

148. May, M. A. and Lumsdaine, A. A. Learning from Films. Yale Univ. Press, New Haven, Conn., 1958.

149. May. M. A. Statistical Methods used in the Character Education Inquiry. In, Hartshorn, H. and May, M. A. Studies on the Organization of Character. New York Macmillan Co. 1930

150. May, M. A. Enchancements and simplifications of motivational and stimulus variables in Audio-Visual instructional materials. USOE (Mimeo.), 109 pp. 1965.

151. May, M. A. Word-picture relationship in Audio-Visual presentation. USOE, (Mimeo.) 115-pp. 1965.

152. McGuigan, F. J., Keller, B. and Stanton, E. Covert language responding during silent reading. J. Educ. Psychol., 1964, 55, 339-343 pp.

153. McGuire, William J. Audience participation and audio-visual instruction: overt-covert responding and rate of presentation. In Lumsdaine, (Ed.), 1961, 417-426 pp.

154. McGuire, William J. Effects of serial position and proximity to "reward" within a demonstration film. In Lumsdaine, (Ed.), 1961, 209-216 pp. (See Ref. 128)

155. McNeil, John D. Programmed instruction as a research tool in reading: an annotated case. Journ. Programmed Instruction, 1, 1, 1962, 37-42 pp.

156. McNeil, J. D. and Keislar, E. R. Questions versus statements as stimuli to children's learning. AV Comm. Rev. 10 1962, 85-88.

157. Mechanic, A. The responses involved in the rote learning of verbal material. J. of Verbal Learning and Verbal Behvior. 3 1964 30-36

158. Mechanic, A. and D'Andrea J. The role of articulation in response learning. Proceedings 73rd Annual Convention of the APA. APA, Washington, D. C., 1965.

159. Mechanic, A. An investigation of selectivity of incidental learners. Proceedings 73rd Annual Convention of the APA. APA, Washington, D. C., 1965.

160. Melarango, R. J. Effect of negative reinforcement in an automated teaching situation. Psychol. Reports, 7, 1960, 381-384.

161. Meyer, Susan R. Report on the initial test of a junior high school vocabulary program. Lumsdaine and Glaser (Eds.), Teaching Machines and Programmed Learning (Washington, D. C.: Department of Audio-Visual Instruction, National Education Association, 1960, 229-46 pp.

162. Michael, Donald N. and Maccoby, Nathan N. Factors influencing the Effects of Student Participation on Verbal Learning From Films. In Lumsdaine (Ed) 1961 271-274 (See Ref. 128)

163. Michael, D. M., Maccoby, Nathan, and Levine, S. Further studies of student participation procedures in film instruction: motivation versus practice effects, "feedback" and overt versus covert responding. In Lumsdaine (ed) 1961. 295-334.

164. Moore, J. William, and Smith, Wendell I. A comparison of several types of "immediate reinforcement". Preliminary report of study supported by Air Force Office of Scientific Research. In Smith and Moore (Eds.), Programmed Learning (New York: Van Nostrand, 1962, 192-201 pp.).

165. Moore, J. William, and Smith, Wendell I. Knowledge of results in self-teaching spelling. Pychol. Reports, 9, 1961, 717-26 pp.

166. Osgood, C. E. Method and Theory in Experimental Psychology. New York, Oxford Press 1953

167. Peterson, H. A. Recitation or recall as a factor in the learning of long prose selections. J. Educ. Psychol. 35, 1944 220-228.

168. Postman, L. Short-term memory and incidental learning. In Melton, A. W. (ed) Categories of Human Learning. New York, Academic Press 1964.

169. Postman, L. and Phillips, L. W. Studies in incidental learning: I. The effects of crowding and isolation. J. Exp. Psychol. 48, 1954, 48-56.

170. Postman, L. and Senders, V. Incidental learning and generality of set. J. Exp. Psychol. 36, 1946, 153-165

171. Premack, D. Toward empirical behavior laws: I. Positive reinforcement. Psychol. Rev. 66 1959, 219-33.

172. Pressey, S. L. Some perspectives and major problems regarding teaching machines. In Lumsdaine and Glaser (eds) Teaching Machines and Programmed Learning, Washington D.C., DAVI, NEA 1960.

173. Pressey, S. L. Development and appraisal of devices providing immediate automatic scoring of objective tests and concomitant self-instruction. In Lumsdaine and Glaser (eds) 1960. (See Ref. 186)

174. Pressey, S. L. Teaching Machine (and learning theory) crisis. J. Applied psychol. 47 1963, 1-6.

175. Rigney, Joseph W., and Budnoff, Irving J. The relative efficiency of different combinations of prompting and confirmation for learning a Boolean algebra program. Prepared for Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research, Los Angeles, Department of Psychology, University of Southern California, 1962. (mimeo.).

176. Ripple, R. E. Comparison of the effectiveness of a programmed text with three other methods of presentation. Psychol. Reports 12, 1963, 227-237

177. Roe, Arnold. Automated teaching methods using linear programs. Jour. Psychol. 40 1962 198-201

178. Roshal, S. M. Film-mediated learning with varying representations of the task: viewing angle, portrayal of demonstration, motion, and student participation. In Lumsdaine (ed.) Student Response in Programmed Instruction, 155-76 (See ref. 128)

179. Rotberg, I. C. and Woolman, M. Verbal paired-associate learning as a function grouping similar stimuli or responses. J. Exp. Psychol. 65, 1961 47-51

180. Rothkopf, E. Z. Automated teaching devices and a comparison of two variations of the method of adjusted learning. Psychol. Rept., 8, 163-69 pp. 1961.

181. Rothkopf, E. Z. Programmed self-instructional booklets, mnemonic phrases and unguided study in the acquisition of equivalences. J. Programmed Instruct., 1 (1), 19-28 pp., 1962.

182. Rothkopf, E. Z. The epidemiology of ignorance. Contemp. Psychol., 9, 234-36, 1964, (Rev. of ref. 141.)

183. Rothkopf, E. Z. and Coke, E. U. Repetition interval and rehearsal method in learning equivalences from written sentences. J. Verbal Learn. and Verbal Behav., 2, 406-16 pp., 1963.

184. Schoer, L. Reactive inhibition as related to performance on programmed learning materials. J. Educ. Psychol. 57 1966, 86-89.

185. Schramm, W. The Research on Programmed Instruction: An annotated bibliography. (Washington, D. C. U.S. Off. Educ., 114 pp., 1964).

186. Sechrest, Lee. Implicit reinforcement of responses. J. Educ. Psychol., 1963, 54, 197-201 pp.

187. Shay, Carleton B. Relationship of intelligence to step size on a teaching machine program. Journ. Educ. Psychol., 52, 1961, 98-103 pp.

188. Sheffield, F. D. Theoretical considerations in the learning of complex sequential tasks from demonstration and practice. In Lumsdaine, A. A. (Ed) Student Response in Programmed Instruction, 1961 31-32 pp.

189. Sheffield, F. D. and Maccoby, N. Summary and interpretation of research on organizational principles in constructing filmed demonstrations. In Lumsdaine, A. A. (Ed) Student Response in Programmed Instruction, 1961 117-31 pp.

190. Sheffield, F. D., Margolius, G. J. and Hoehn, A. J. Experiments on perceptual mediation in the learning of organizable sequences. In Lumsdaine, A. A., (Ed.), Student Response in Programmed Instruction, 1961 107-16 pp.

191. Shettel, H. H., Angell, D., Fuzio, J. L., Lindley, R. H. and Lumsdaine, A. A. Training Requirements for the Development of Prototype Self-Instructional Programs Covering Basic Job Knowledge for SAGE Track Monitor and Intercept Director Operator Positions (Am. Inst. for Res. Rept., No. AIR-C11-61-SR-246, Pittsburgh, Pa., 34 pp., March, 1961).

192. Shettel, Harris, H. and Lindley, Richard H. An experimental comparison of two types of self-instructional programs for a SAGE system paired-associate task. Pittsburgh, Pa.: American Institute for Research, 1961, 21pp. (offset).

193. Sidowski, J., Kopstein, F. F. and Shillestad, I. J. Prompting and confirmation variables in verbal learning. Psychol. Rept., 8, 401-6, 1961.

194. Silbeman, H. F. Self-teaching devices and programmed materials. Rev. of Educational Research 1962, 32 (2) 179-193.

195. Silberman, H. F. and Carter, L. F. The Systems Approach, Technology and the School. In New Approaches to Individualized Instruction. Princeton, N. J. Educational Testing Service 1965.

196. Silberman, H. F. et al. Fixed sequence versus branching auto-instructional methods. J. Educ. Psychol., 62, 166-72 pp., 1961.

197. Silberman, H. F. et al., Development and Evaluation of Self-Instructional Materials for Underachieving and Overachieving Students (System Developm. Corp., Santa Monica, Calif., 115 pp., July 1962).

198. Silberman, Harry F., Melaragno, Ralph J. and Coulson, John E. Confirmation and prompting with connected discourse material. Psychol. Reports, 9, 1961, 235-38 pp.

199. Silverman, R. E. The Comparative Effectiveness of Animated and Static Transparencies (U. S. Naval Train. Device Cent., ONR. Tech. Rept. No. SDC 78-1, Port Washington, N. Y., April 1958.)

200. Silverman, Robert E. and Alter, Millicent. Note on the response in teaching machine programs. Psychol. Reports, 7, 1960, 496 pp.

201. Skinner, B. F. Teaching machines. Science, 128, 969-77 (1958) Reprint in Teaching Machines and Programmed Learning, 137-58.

202. Skinner, B. F. The programming of verbal knowledge. In Automatic Teaching: The State of the Art. 63-68 (Galanter, E.H., Ed., Wiley, New York, 198 pp., 1959).

203. Skinner, B. F. Why we need teaching machines. Harv. Educ. Rev., 31, 377-98, 1961.

204. Skinner, B. F. Reflections on a decade of teaching machines. In Glaser (ed) Teaching Machines and Programmed Instruction II. Washington, D. C. DAVI, NEA 1965.

205. Smith, Phillip D. Knowledge of Results and Continuity of Various Techniques in Presenting a Filmstrip as Factors in Immediate Learning and Retention. Greenville, C.C.: School of Education, Bob Jones University 1963, 234p

206. Smith, Wendell, and Moore, J. William. Size of step and achievement in programmed spelling. Lewisburg, Pa.: Bucknell University, 1961, 15 pp.

207. Smith, Wendell, I. and Moore, J. William. Size of step and achievement in programmed spelling. Psychol. Reports, 10 (1962), 287-94. (Also reprinted in Smith and Moore, Programmed Learning, New York: Van Nostrand, 1962, 202-06 pp.)

208. Stevens, L. F. and Sherman, M. A. Eye fixations as a function of response contingent material in a teaching machine program. Paper read at Eastern Psychological Association Annual convention 1965, page 116 of Abstracts.

209. Stolurow, L. M. and Walker, C. C. A comparison of overt and covert response in programmed learning. Journ. Educ. Research, 55, 1962, 421-32 pp.

210. Stolurow, Lawrence M. A comparative study of methods of programming materials for efficient learning in self-instructional devices. Urbana, Illinois: University of Illinois, 1961, 14 pp. (mimeo.).

211. Stolurow, Lawrence M. Implications of current research for future trends. Journ. Educ. Res., 1962, 55, 519-527 pp.

212. Stolurow, L. M. and Davis, D. Teaching Machines and computer-based instruction. In. Glaser, R. (ed) Teaching Machines and Programmed Learning II. Washington, D. C. DAVI, NEA 1965.

213. Suppes, Patrick and Ginsberg, Rose. Application of a stimulus sampling model to children's concept formation with and without overt correction response. Journ. Exp. Psychol., 63, 1962, 330-36 pp.

214. Taber, J. I. and Glaser, R. An exploratory evaluation of discrimination transfer learning using literal prompts. J. Educ. Res. 55 1962, 508-512.

215. Taber, J. I., Glaser, R., and Schaefer, H. H. Learning and Programmed Instruction. Reading, Mass. Addison-Wesley, 1965.

216. Taylor, J. A. and Spence K. W. The relationship of anxiety level to performance in serial learning. J. Exp. Psychol. 44 1952, 61-64

217. Thomson, L. The Role of Verbalization in Learning From Demonstrations. Unpublished Ph.D. dissertation, Yale University 1944.

218. Thorndike, E. L. Ideo-motor action. Psychol. Rev. 20 1913, 91-105

219. Tobias, S. and Weiner, M. Effect of response modes in a review program. J. Programmed Instruction 2, 1963, 9-18.

220. Travers, R. M. W. Essentials of Learning: An Overview for Students of Education. New York, Macmillan Co. 1963, 544 p.

221. Travers, R. M. W., Keith, R. VanWagenen, D. H. and McCormick, Mary. Learning as a consequence of the learner's task involvement under different conditions of feedback. J. Educ. Psychol. 55 1964, 167-173.

222. Underwood, B. J. Verbal learning in the educative process. Harvard Educational Review *29* 1959, 107-117.

223. Underwood, B. J. and Schultz, R. W. *Meaningful Verbal Learning*. Chicago, J. B. Lippencott Co. 1960.

224. Van Wagenen, R. K. and Travers, R. M. W. Learning under conditions of direct and vicarious reinforcement. *J. Educ. Psycho.*, 1953, *54*, 347-362 pp.

225. Vicory, A. C. *Preliminary Research Report of the Corrigan Telecommunication System*. San Jose, Calif. 1963.

226. Vicory, A. C. and Corrigan, Robert E. Learning math concepts through required responses and feedback. San Jose, California: Instructional Television, San Jose State College, 1963, 24 pp. (mimeo.).

227. Wallace, John. Concept dominance, type of feedback and intensity of feedback as related to concept attainment. *J. Educ. Psychol.*, 1964, *55*, 159-166 pp.

228. Wallace, W. P. and Underwood, Benton. Implicit responses and the role of intralist similarity in verbal learning by normal and retarded students. *J. Educ. Psychol.*, 1964, *55*, 362-370 pp.

229. Weiss, Walter, Maccoby, Nathan and Sheffield, Fred. D. Combining practice with demonstration in teaching complex sequences: serial learning of a geometric construction task. In Lumsdaine (Ed.), 1961, 55-76 pp.

230. Williams, J. P. Comparison of several response modes in a review program. J. Educ. Psychol., 54, 253-60, 1963.

231. Wittrock, M. C. Programming portrayed in breadth - not depth. Contemp. Psychol., 8, 172-73 pp., 1963, (Rev. of ref. 170).

232. Wittrock, M. C. Response mode in the programming of kinetic molecular theory concepts. J. Educ. Psychol., 54, 1963, 89-93 pp.

233. Wittrock, M. C. and Twelker, P. A. Prompting and feedback in the learning retention, and transfer of concepts. Brit J. Educ. Psychol., 34, 10-18 pp., 1964.

234. Wolgamuth, D. A. Comparative study of three techniques of student feedback in television teaching: The effectiveness of an electrical signal feedback system. (American Univ., USOE Proj. No. 453, Washington, 75 pp., December 1961) (Abstract, in AV Commun. Rev., USOE Install. 4, 100-1).

4/

235. Wulff, J. J. and Emeson, D. L. The relationship between "what is learned" and "how it is taught". In Lumsdaine, A. A. (Ed) Student Response in Programmed Instruction, 1961, 457-70 pp.

236. Wulff, J. J. and Kraeling, D. G. Familiarization procedures used as adjuncts to assembly-task training with a demonstration film. In Lumsdaine, A. A. (Ed) Student Response in Programmed Instruction, 1961 141-53 pp.

237. Wulff, J. J. and Stolurow, L. M. The role of class-descriptive cues in paired-associate learning. In Lumsdaine, A. A. (Ed) Student Response in Programmed Instruction, 1961 445-55 pp.

238. Wynn, D. R. and McKeegan H. F. Assessment of a Graduate Level Self-Teaching Technique Based on a Combination of Programmed Instruction and the Case Method. Pittsburgh, Pa. Univ. of Pittsburgh, School of Education, 1963, 108 p. (University Microfilms Pub. No 64-13951)

239. Zeaman, David. Skinner's theory of teaching machines. In: Salanter, Eugene (Ed.)., Automatic Teaching, the state of the art. New York: Wiley, 1954.

240. Zuckerman, C. B., Marshall, G. R., and Groesberg, S. Research in the Automation of Teaching. Port Washington, N. Y. U. S. Naval Training Devices Center, 1961